

doesn't come up on applicant

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:36:41 ; Search time 25.93 Seconds  
(without alignments)  
21.099 Million cell updates/sec

Title: US-08-765-837-5  
Perfect score: 86  
Sequence: 1 ENRQGAEPKRNFKVP 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 112136

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_36.\*  
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	69.8	15	18 W08420	Factor IX binding
2	36	41.9	16	18 W08419	Factor IX binding
3	33	38.4	15	18 W08421	Factor IX binding
4	30	34.9	8	19 W75817	Mouse mast cell pr
5	30	34.9	14	16 R83129	Human GMPS tryptic
6	28	32.6	8	16 R66094	Myelopietic proge
7	28	32.6	12	18 W34873	Human tau protein
8	28	32.6	13	20 Y42676	HHV-6 variant A de
9	28	32.6	14	16 R66092	Myelopietic proge
10	28	32.6	14	19 Y20167	Human P53 immunog
11	28	32.6	14	19 W81782	Human H2A epitope
12	28	32.6	15	12 R12025	N-terminal of prot

13	28	32.6	15	19 W79210	Framework 2 region
14	27	31.4	13	17 R95642	Antigen peptide fr
15	27	31.4	14	18 W01827	CD44 epitope. Rat
16	27	31.4	15	15 R55685	55-63 kl. glial gro
17	27	31.4	16	21 Y59337	Tyrosine phosphata
18	26	30.2	9	20 W99342	Mouse IgG2A upper
19	26	30.2	10	20 Y46552	Immunogenic peptid
20	26	30.2	11	18 W44211	Human soluble RAGE
21	26	30.2	11	18 W33765	Human RAGE polypep
22	26	30.2	12	21 Y91911	Peptide 23 derived
23	26	30.2	13	17 R91945	TNF receptor bindi
24	26	30.2	13	19 W69499	Fragment of G. oxy
25	26	30.2	14	15 R55176	Mouse IgG1 hinge s
26	26	30.2	15	12 R10061	Growth hormone rel
27	26	30.2	15	14 R37640	Sequence of IgG2A
28	26	30.2	15	17 Y16424	Peptide fragment o
29	26	30.2	15	20 Y13338	Naturally occurin
30	26	30.2	15	20 W93125	Human growth hormo
31	26	30.2	15	20 W88090	C-terminal hinge r
32	26	30.2	15	21 Y98918	HLA class II bindi
33	26	30.2	15	21 Y57245	IgG2A mouse antige
34	26	30.2	15	21 Y80915	Murine IgG2A hinge
35	26	30.2	16	13 R20012	N-terminal sequenc
36	26	30.2	16	14 R31436	Growth hormone rel
37	26	30.2	16	15 R82023	Stearoyl-ACP-desat
38	25	29.1	9	15 R57579	Peptide (79) inhib
39	25	29.1	9	15 R57722	Peptide (222) inhi
40	25	29.1	9	15 R57651	Peptide (151) inhi
41	25	29.1	10	11 R03909	HIV-antibody react
42	25	29.1	11	21 Y88537	NCAM Igl binding p
43	25	29.1	13	13 R21059	Sequence of peptid
44	25	29.1	13	16 R82711	Shrimp tropomyosin
45	25	29.1	13	18 W18067	Soluble starch syn

ALIGNMENTS

RESULT 1  
W08420  
ID W08420 standard; peptide: 15 AA.  
XX  
AC W08420;  
XX  
DT 04-SEP-1997 (first entry)  
XX  
DE Factor IX binding peptide #4 from factor VIII A3 domain exosite 1.  
XX  
KW Exosite: factor VIII; A3 domain; activated factor IX; precursor;  
KW thrombotic disorder; inhibition; intrinsic blood coagulation pathway.  
XX  
OS Homo sapiens.  
XX  
FN W09641816-A1.  
XX  
PD 27-DEC-1996.  
XX  
PF 12-JUN-1996; 96WO-NL00236.  
XX  
PR 12-JUN-1995; 95EP-0201554.  
XX  
PA (BLOE-) STICHTING CENT LAB VAN DE BLOEDTRANSFUSI.  
XX  
PI Lenting PJ, Mertens K;  
XX  
DR WPI; 1997-065422/06.  
XX  
PT Factor VIII peptide that binds to factor IX - useful for treating  
PT thrombotic disorders  
XX  
PS Example 2; Page 17; 39pp; English.  
XX  
CC The sequences given in W08417-24 represent peptides derived from

CC exosite I derived from factor VIII A3 domain. The exosite regions  
 CC have low hydrophathy values which reflects their hydrophilic nature.  
 CC These peptides and fragments derived from them are capable of binding-  
 CC activated factor IX and its non-activated precursor. These peptides  
 CC can be used for treating thrombotic disorders by inhibiting the  
 CC intrinsic blood coagulation pathway. This peptide corresponds to  
 CC Factor VIII A3 domain region 1799-1813.  
 CC  
 XX Sequence 15 AA;

SQ

Query Match 69.8%; Score 60; DB 18; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00024;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAEPKRNFKVP 16  
 DB 1 gaepkrnfkvp 11  
 |||||

RESULT 2

W08419  
 ID W08419 standard; peptide: 16 AA.

AC W08419;

DT 04-SEP-1997 (first entry)

DE Factor IX binding peptide #3 from factor VIII A3 domain exosite I.

EX Exosite; factor VIII; A3 domain; activated factor IX; precursor;

KW thrombotic disorder; inhibition; intrinsic blood coagulation pathway.

XX Homo sapiens.

OS

XX W09641816-A1.

PN 27-DEC-1996.

PD 12-JUN-1996; 96WO-NL00236.

PF 12-JUN-1995; 95EP-0201554.

PR (BLOE-) STICHTING CENT LAB VAN DE BLOEDTRANSFUSI.

PA Lenting PJ, Mertens K;

PI WPI; 1997-065422/06.

DR Factor VIII peptide that binds to factor IX - useful for treating

XX thrombotic disorders

PS Example 2; Page 17; 39pp; English.

XX The sequences given in W08417-24 represent peptides derived from  
 CC exosite I derived from factor VIII A3 domain. The exosite regions  
 CC have low hydrophathy values which reflects their hydrophilic nature.  
 CC These peptides and fragments derived from them are capable of binding-  
 CC activated factor IX and its non-activated precursor. These peptides  
 CC can be used for treating thrombotic disorders by inhibiting the  
 CC intrinsic blood coagulation pathway. This peptide corresponds to  
 CC Factor VIII A3 domain region 1786-1801.  
 CC  
 XX Sequence 16 AA;

SQ

Query Match 41.9%; Score 36; DB 18; Length 16;

Best Local Similarity 87.5%; Pred. No. 4.6;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENROQGAE 8

DB 9 edrqgae 16

|||||

RESULT 3

W08421  
 ID W08421 standard; peptide: 15 AA.

XX W08421;

AC

DT 04-SEP-1997 (first entry)

DE Factor IX binding peptide #5 from factor VIII A3 domain exosite I.

EX Exosite; factor VIII; A3 domain; activated factor IX; precursor;

KW thrombotic disorder; inhibition; intrinsic blood coagulation pathway.

XX Homo sapiens.

OS W09641816-A1.

PN 27-DEC-1996.

PD 12-JUN-1996; 96WO-NL00236.

PF 12-JUN-1995; 95EP-0201554.

PR (BLOE-) STICHTING CENT LAB VAN DE BLOEDTRANSFUSI.

PA Lenting PJ, Mertens K;

PI WPI; 1997-065422/06.

DR Factor VIII peptide that binds to factor IX - useful for treating

XX thrombotic disorders

PS Example 2; Page 17; 39pp; English.

XX The sequences given in W08417-24 represent peptides derived from

CC exosite I derived from factor VIII A3 domain. The exosite regions

CC have low hydrophathy values which reflects their hydrophilic nature.

CC These peptides and fragments derived from them are capable of binding-

CC activated factor IX and its non-activated precursor. These peptides

CC can be used for treating thrombotic disorders by inhibiting the

CC intrinsic blood coagulation pathway. This peptide corresponds to

CC Factor VIII A3 domain region 1804-1818.

XX Sequence 15 AA;

SQ

Query Match 38.4%; Score 33; DB 18; Length 15;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 KNFVKP 16

DB 1 knfvkp 6

|||||

RESULT 4

W75817

ID W75817 standard; peptide: 8 AA.

XX W75817;

AC

DT 27-OCT-1998 (first entry)

DE Mouse mast cell protease (mMCP-6) susceptible peptide 11.

EX Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;

KW trypsin-6 protein; inflammatory disorder; allergic rhinitis; urticaria;

KW antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;

KW hyperproliferative skin disease; peptic ulcer; hypersensitivity;

KW inflammatory skin condition.

XX

OS Synthetic.  
 OS Mus sp.  
 PN WO9833812-A1.  
 XX  
 PD 06-AUG-1998.  
 XX  
 PF 30-JAN-1998; 98WO-US01865.  
 XX  
 PR 05-FEB-1997; 97US-0037090.  
 XX  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 XX  
 PI Huang C, Stevens RL;  
 XX  
 DR WPI; 1998-437390/37.  
 XX  
 PT Trypsinase-6 complex inhibitory peptides - used to treat mast  
 PT cell-mediated inflammatory disorders e.g. asthma  
 XX  
 PS Examples; Page 26; 69pp; English.  
 XX  
 CC Sequences shown in W75807 to W75835 represent mouse mast cell protease  
 CC (mMCP-6) susceptible peptides obtained in the absence of heparin. The  
 CC invention provides sequences shown in W63160 to W63169 that are  
 CC inhibitors of mMCP-6. These trypsinase-6 complex inhibitor peptides can  
 CC be used for treating a mast cell-mediated inflammatory disorder. The  
 CC inhibitors can be used to treat inflammatory disorders including asthma,  
 CC allergic rhinitis, urticaria and antioedema, eczematous dermatitis  
 CC (atopic dermatitis), hyperproliferative skin disease, anaphylaxis, peptic  
 CC ulcers, inflammatory bowel disorder, hyperresponsiveness and inflammatory  
 CC skin conditions.  
 XX  
 SQ Sequence 8 AA;

Query Match 34.9%; Score 30; DB 19; Length 8;  
 Best Local Similarity 62.5%; Pred. No. 2,1e+05;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QROGAEPK 10  
 I::I I::I  
 Db 1 qkrgrepr 8

RESULT 5  
 R83129  
 ID R83129 standard; Peptide; 14 AA.  
 XX  
 AC R83129;  
 XX  
 DT 31-JAN-1996 (first entry)  
 XX  
 DE Human GMPS tryptic peptide # 6.  
 XX  
 KW Human guanosine 5'-monophosphate synthetase; A3.01 cells;  
 KW tryptic peptide.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9527789-A.  
 XX  
 PD 19-OCT-1995.  
 XX  
 PF 07-APR-1995; 95WO-US03934.  
 XX  
 PR 08-APR-1994; 94US-0224917.  
 XX  
 PA (SYNT ) SYNTEX USA INC.  
 XX  
 PI Barnett JW, Lou L;  
 XX  
 DR WPI; 1995-366393/47.

XX  
 PT New isolated human guanosine 5'-mono-phosphate synthetase - used to  
 PT develop prods. for its study and for identifying inhibitors useful for  
 PT e.g. anti-cancer or immunosuppressive therapy  
 XX  
 PS Example; Table 2, page 18; 48pp; English.  
 XX  
 CC Naturally occurring human GMPS was purified from A3.01 cells and  
 CC digested with trypsin. Nine tryptic peptides were resolved. Their  
 CC sequences are given in R83124-R83132 and are indicated on R83122 PW.  
 CC Based on the peptide sequences, degenerate oligos were synthesised  
 CC in both the sense and antisense orientations and used in PCR. A  
 CC fragment was generated with oligos 2S2 and 8A2 (see T00493 FT). 2S2  
 CC & 8A2 corresp. to tryptic peptides 2 & 8. This PCR fragment (PCR  
 CC 2S8A) (see T00493 FT) was used to screen an A3.01 cDNA library. The  
 CC complete sequence of positive clone 6 (GMPS.6 T00492) was determined  
 CC and is shown in Figure 1 (T00493). The derived AA sequence (R83123)  
 CC of human GMP synthetase is shown in Figure 1 (R83122). The  
 CC predicted mol. wt. of the enzyme - 76,725 - was in good agreement  
 CC with the size indicated by polyacrylamide gel electrophoresis of  
 CC the purified A3.01 human GMPS.  
 XX  
 SQ Sequence 14 AA;

Query Match 34.9%; Score 30; DB 16; Length 14;  
 Best Local Similarity 62.5%; Pred. No. 46;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 EPRKNFVK 15  
 I I I I I I  
 Db 3 eplkfik 10

RESULT 6  
 R66094  
 ID R66094 standard; peptide; 8 AA.  
 XX  
 AC R66094;  
 XX  
 DT 12-AUG-1995 (first entry)  
 XX  
 DE Myelopoietic progenitor cell inhibitor peptide.  
 XX  
 KW Myelopoietic; myeloid; progenitor cell inhibitor; leukaemia;  
 KW polycythemia; myelosuppression; septic shock; hypotension.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9428013-A.  
 XX  
 PD 08-DEC-1994.  
 XX  
 PF 18-MAY-1994; 94WO-US05773.  
 XX  
 PR 20-MAY-1993; 93US-0065722.  
 XX  
 PA (RESE ) RESEARCH CORP TECHNOLOGIES INC.  
 XX  
 PI Broxmeyer HE, Cooper S, Kreisberg M, Kreisberg R;  
 PI Lu L, Moore RN;  
 XX  
 DR WPI; 1995-022707/03.  
 XX  
 PT New peptide(s) that suppress myeloid progenitor cell  
 PT proliferation - contain the sequence Ala Lys Pro Arg, useful for  
 PT treating or preventing e.g. leukaemia, radiation induced  
 PT myelosuppression and septic shock  
 XX  
 PS Claim 5; Page 40; 50pp; English.  
 XX  
 CC New peptides are claimed which contain at least 5 (pref. 5-20) amino  
 CC acids and which include the sequence Ala-Lys-Pro-Arg. The present

CC sequence is a specifically claimed such peptide.  
 CC The peptides suppress proliferation of myeloid progenitor  
 CC cells and reduce susceptibility to septic shock. They are used to  
 CC treat chronic myelogenous leukemia or polycythemia, to reduce  
 CC chemotherapy- or radiation-induced myelosuppression, to prevent  
 CC septic shock and to reduce nitric oxide-induced hypotension.  
 XX  
 SQ Sequence 8 AA;

Query Match 32.6%; Score 28; DB 16; Length 8;  
 Best Local Similarity 71.4%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 AEPRKNF 13  
 |||||  
 Db 1 akpranf 7

RESULT 7  
 W34873  
 ID W34873 standard; peptide; 12 AA.  
 XX W34873;  
 AC  
 XX 27-MAR-1998 (first entry)  
 DT  
 XX Human tau protein fragment.  
 DE  
 XX  
 XX Antibody: phosphorylated tau protein; paired helical filament;  
 KW detection; Alzheimer's disease; human.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO9734145-A1.  
 PN  
 XX  
 XX 18-SEP-1997.  
 PD  
 XX  
 XX 13-MAR-1997; 97WO-JP00804.  
 PF  
 XX  
 XX 13-MAR-1996; 96JP-0056090.  
 PR  
 XX (MITU ) MITSUBISHI CHEM CORP.  
 PA  
 XX Imahori K, Ishiguro K, Park J, Sato K, Uchida T;  
 PI WPI; 1997-470978/43.  
 DR  
 XX  
 XX Antibody prepared using a partial peptide containing part of  
 PT phosphorylated tau protein - used for detecting Alzheimer's disease  
 PS Example; Page 36; 48pp; Japanese.  
 XX  
 XX An antibody, prepared using a partial peptide containing the  
 CC phosphorylated residue of the phosphorylated tau protein, e.g. the  
 CC present sequence, in a paired helical filament, can be used to  
 CC detect Alzheimer's disease, i.e. by detecting phosphorylated tau  
 CC protein in brain extracts or tissue fragments.  
 XX  
 SQ Sequence 12 AA;

Query Match 32.6%; Score 28; DB 18; Length 12;  
 Best Local Similarity 71.4%; Pred. No. 89;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 AEPRKNF 13  
 |||||  
 Db 1 aeprqef 7

RESULT 8  
 Y42676

ID Y42676 standard; peptide; 13 AA.  
 XX  
 AC Y42676;  
 XX  
 DT 17-JAN-2000 (first entry)  
 XX  
 DE HHV-6 variant A derived peptide epitope.  
 XX  
 XX Human herpes virus-6; HHV-6; infection; immunological; MIEP; epitope;  
 KW major immediate early protein.  
 XX  
 OS Synthetic.  
 OS Human herpesvirus 6.  
 XX  
 PN WO9949086-A1.  
 XX  
 PD 30-SEP-1999.  
 XX  
 XX 26-MAR-1999; 99WO-US06921.  
 PF  
 XX 26-MAR-1998; 98US-0079379.  
 PR  
 XX 23-MAR-1999; 99US-0274938.  
 XX

PA (CARR/) CARRIGAN D R.  
 PA (KEHL/) KEHL K K.  
 XX  
 XX Carrigan DR, Kehl KK;  
 PI  
 XX  
 XX WPI; 1999-601224/51.  
 DR

XX Methods for the rapid detection of human herpes virus 6 variants A and  
 PT B utilizing antibodies raised against synthetic peptides  
 XX  
 XX Claim 15; Page 42; 58pp; English.  
 XX

CC The invention relates to a method for detecting human herpes virus  
 CC (HHV)-6 infection that comprises contacting host cells with immunological  
 CC reagents specific for an epitope of HHV-6 variant A and/or B major  
 CC immediate early protein (MIEP). HHV-6 peptides for raising an  
 CC immunological reagent that binds specifically to an epitope of: (a)  
 CC HHV-6 variant A and not B MIEP; or (b) HHV-6 variant A and B MIEP are  
 CC also provided. The peptides are useful as immunological reagents, e.g.  
 CC hyperimmune sera. Monoclonal antibodies and recombinant DNA-derived  
 CC single chain fragment variables (ScFv) may be useful for detecting  
 CC HHV-6 infection when raised against specific epitopes of the HHV-6  
 CC MIEP. The methods provide a rapid culture procedure having a high level  
 CC of sensitivity and specificity. The methods also have shortened  
 CC turnaround time (compared to other known methods, e.g. the isolation of  
 CC the virus in cell culture, the detection of virus DNA in an acellular  
 CC specimen by PCR or positive immunohistochemical staining) and can be  
 CC easily implemented by a diagnostic laboratory.  
 XX  
 XX Sequence 13 AA;  
 SQ

Query Match 32.6%; Score 28; DB 20; Length 13;  
 Best Local Similarity 41.7%; Pred. No. 96;  
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENORQGAEPKRN 12  
 ||||| : |  
 Db 2 ensrqsgdeqtn 13

RESULT 9  
 R66092  
 ID R66092 standard; peptide; 14 AA.  
 XX  
 AC R66092;  
 XX  
 DT 12-AUG-1995 (first entry)  
 XX  
 XX Myelopoietic progenitor cell inhibitor peptide.



```

XX Myelopietic; myeloid; progenitor cell inhibitor; leukaemia;
KW Polycythemia; myelosuppression; septic shock; hypotension.
XX Synthetic.
XX OS
XX WO9428013-A.
PN
XX
XX 08-DEC-1994.
PD
XX
XX 18-MAY-1994; 94WO-US05773.
PF
XX
XX 20-MAY-1993; 93US-0065722.
PR
XX
XX (RESE ) RESEARCH CORP TECHNOLOGIES INC.
PA
XX
XX Broxmeyer HE, Cooper S, Kreisberg M, Kreisberg R;
PI Lu L, Moore RN;
PI
XX
XX WPI; 1995-022707/03.
DR
XX
XX New peptide(s) that suppress myeloid progenitor cell
PT proliferation - contain the sequence Ala-Lys-Pro-Arg, useful for
PT treating or preventing e.g. leukaemia, radiation induced
PT myelosuppression and septic shock
PT
XX
XX Claim 3; Page 40; 50pp; English.
PS
XX
XX New peptides are claimed which contain at least 5 (pref. 5-20) amino
CC acids and which include the sequence Ala-Lys-Pro-Arg. The present
CC sequence is a specifically claimed such peptide.
CC The peptides suppress proliferation of myeloid progenitor
CC cells and reduce susceptibility to septic shock. They are used to
CC treat chronic myelogenous leukemia or polycythemia, to reduce
CC chemotherapy- or radiation-induced myelosuppression, to prevent
CC septic shock and to reduce nitric oxide-induced hypotension.
CC The present sequence is the preferred peptide for clinical use.
XX
XX Sequence 14 AA;
SQ
Query Match 32.6%; Score 28; DB 16; Length 14;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 7 ABRKNF 13
Db 7 akpranf 13
RESULT 10
Y20167
ID Y20167 standard; Peptide; 14 AA.
XX
XX Y20167;
AC
XX
XX 22-JUL-1999 (first entry)
DT
XX
XX Human P53 immunogenic peptide fragment.
DE
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A;
KW immunogenic.
XX
XX Homo sapiens.
PN
XX OS

```

---

```

XX WO9845322-A2.
PN
XX
XX 15-OCT-1998.
PD
XX
XX 02-APR-1998; 98WO-IB00705.
PF
XX
XX 10-APR-1997; 97US-0043163.
PR
XX
XX (UYUT-) RIJKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
PI
XX
XX WPI; 1998-609901/51.
DR
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX
XX Claim 24; Page 85; 258pp; English.
PS
XX
XX This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX Sequence 14 AA;
SQ
Query Match 32.6%; Score 28; DB 19; Length 14;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 QROGAEPR 10
Db 1 qergaspr 8
RESULT 11
W81782
ID W81782 standard; Protein; 14 AA.
XX
XX W81782;
AC
XX
XX 23-FEB-1999 (first entry)
DT
XX
XX Human HE4 epitope peptide #2.
DE
XX
XX HE4; epidiidymis-specific; diagnosis; male infertility; treatment;
KW sterility; immunosterilisation.
KW
XX
XX Homo sapiens.
OS
XX
XX EP878544-A1.
PN
XX OS

```

```
PD 18-NOV-1998.
XX
PF 29-JAN-1991; 91EP-0250021.
XX
PR 30-NOV-1990; 90DE-4038189.
XX
PR 01-FEB-1990; 90DE-4002981.
XX
PA (IHFH-) IHF INST HORMON & FORTPLANZUNGS.
XX
XX Ivell R, Kirchhoff C;
XX WPI; 1998-585748/50.
XX
XX DNA encoding human epididymis polypeptides - useful for, e.g.
PT diagnosis of male infertility
XX
XX Example 9; Page 19; 29pp; German.
XX
XX This sequence represents a novel human epididymis-specific protein, HE4
CC epitope fragment. This protein may be used for cloning and for expression
CC of human epididymis-specific polypeptides in prokaryotic or eukaryotic
CC host cells. Such proteins and antibodies generated from them may be used
CC for diagnosis of e.g. male infertility. The polypeptides and antibodies
CC may also be used for treatment of male infertility and for
CC immunosterilisation of mammals.
XX
XX Sequence 14 AA;
SQ

Query Match 32.6%; Score 28; DB 19; Length 14;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 NOROGAEPRKN 12
Db |::|: |::|
1 ndkegsapqvn 11

RESULT 12
ID R12025 standard; Protein; 15 AA.
XX
AC R12025;
XX
DT 27-NOV-1991 (first entry)
XX
DE N-terminal of protein C analogue.
XX
KW Protein C.
XX
OS Synthetic.
XX
XX JP03091479-A.
XX
XX 17-APR-1991.
XX
XX 05-SEP-1989; 89JP-0229539.
XX
XX 05-SEP-1989; 89JP-0229539.
XX
XX (TEIJ ) TEIJIN KK.
XX
XX WPI; 1991-276411/38.
XX
XX N-PSDB; Q11737.
XX
XX Protein having human protein C and activated human protein C
PT activity - has natural human protein structure with aminoacid
PT from 12th to terminal position e.g. aspartic acid.
XX
XX Claim 1; Page 1; 14pp; Japanese.
XX
XX The sequence is one of several new N-terminals for human
XX protein C. see also R12022-R12031.
CC

XX
SQ Sequence 15 AA;

Query Match 32.6%; Score 28; DB 19; Length 15;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ENORGAEPK 11
Db |::|: |::|
3 edgedqvprk 13

RESULT 13
ID W79210 standard; Protein; 15 AA.
XX
AC W79210;
XX
DT 21-DEC-1998 (first entry)
XX
DE Framework 2 region of human V kappa gene HUM5400.
XX
KW Monoclonal antibody; MAB; LO-CD2a; humanised antibody; CD2 antigen;
KW human lymphocyte; immune response; chimeric; graft-versus-host disease;
KW T-cell; transplant rejection; autoimmune disease; HUM5400.
XX
XX Homo sapiens.
XX
XX US5817311-A.
XX
XX 06-OCT-1998.
XX
XX 07-JUN-1995; 95US-0472281.
XX
XX 07-JUN-1995; 95US-0472281.
XX
XX 05-MAR-1993; 93US-0027008.
XX
XX 09-SEP-1993; 93US-0119032.
XX
XX 29-MAR-1995; 95US-0407009.
XX
XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
XX Bazin H, Latinne D;
XX
XX WPI; 1998-556337/47.
XX
XX Inhibition of T-cell mediated immune response with anti-CD2
PT monoclonal antibody LO-CD2a - used for preventing transplant
PT rejection or for treating graft-versus-host disease or auto-immune
PT diseases
XX
XX Example 7; Columns 33-34; 96pp; English.
XX
XX This represents the amino acid sequence of the framework 2 region of
CC human V kappa gene HUM5400. This is used to construct a humanised antibody
CC LO-CD2a. The invention relates to the use of the monoclonal antibody
CC (MAB) LO-CD2a or a humanised or a chimeric version of the LO-CD2a
CC antibody for the inhibition of a T-cell mediated immune response in a
CC patient. The MAB LO-CD2a (produced by hybridoma cell line AHC MB 11423)
CC can bind to an epitope on the CD2 antigen of the human lymphocytes. The
CC T-cell mediated immune response in a patient can be inhibited by
CC administering the MAB LO-CD2a or an antibody that binds to the same
CC human lymphocyte epitope as LO-CD2a. The method is used for preventing
CC transplant rejection or for treating graft-versus-host disease or for
CC treating autoimmune diseases.
XX
XX Sequence 15 AA;
SQ

Query Match 32.6%; Score 28; DB 19; Length 15;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

QY \*3 QROGAEPK 11  
 || | ||:  
 Db 4 qrpqgspr 12

## RESULT 14

R95642  
 ID R95642 standard; Peptide; 13 AA.

XX AC R95642;

XX DT 26-OCT-1996 (first entry)

XX DE Antigen peptide from cartilage-derived morphogenetic protein-1.

XX KW Human; antigen; cartilage-derived morphogenetic protein-1; CDMP-1;  
 KW polyclonal antibody; rabbit; articular cartilage; chondrogenic;  
 KW vulnery; implantation; chondromalacia; osteoarthritis; therapy;  
 KW joint repair.

XX OS Homo sapiens.

XX PN W09614335-A1.

XX PD 17-MAY-1996.

XX PF 07-NOV-1994; 94WO-US12814.

XX PR 07-NOV-1994; 94WO-US12814.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Chang SC, Luyten FP, Moos M;

XX DR WPI; 1996-251714/25.

XX PT New purified cartilage extracts and proteins - used to stimulate the  
 XX development and repair of cartilage in vivo.

XX PS Example 7; Page 15; 34pp; English.

XX CC This peptide antigen is derived from the mature C-terminal domain  
 CC of human articular cartilage-derived morphogenetic protein-1 (CDMP-1,  
 CC R95635), and has been used to generate polyclonal antibodies in  
 CC rabbits. The peptide does not show sequence identity with any  
 CC other known protein, including bone morphogenetic protein. The  
 CC resulting antibodies have been used to screen tissues from human  
 CC embryos, to study differential CDMP gene expression. At 6 wk,  
 CC CDMP-1 is detected in precartilaginous condensations, and at 7.5-8.5 wk  
 CC CDMP-1 is found in cartilaginous cores of long bones. In areas of  
 CC active cartilage degradation and bone matrix formation, CDMP-1  
 CC expression is also detected in hypertrophic chondrocytes. No  
 CC expression is detected in the axial skeleton, and only low levels  
 CC are present in other tissues. CDMP-1 is present in a purified  
 CC cartilage extract (claimed) which stimulates local cartilage  
 CC formation and repair when combined with a matrix and implanted  
 CC in a mammal. The protein may be used in therapy of e.g.  
 CC chondromalacia or osteoarthritis, to heal joint surfaces, or to  
 CC repair cartilage after reconstructive surgery.

XX SQ Sequence 13 AA;

Query Match 31.4%; Score 27; DB 17; Length 13;  
 Best Local Similarity 62.5%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 QGAEPKRN 12

Db 1 qgkrpskn 8

## RESULT 15

W01827

ID W01827 standard; peptide; 14 AA.

XX AC W01827;

XX DT 20-JUL-1997 (first entry)

XX DE CD44 epitope.

XX KW T-cell; receptor; CD44; epitope; fusion protein; cancer;  
 KW immunoglobulin.

XX OS Rattus rattus.

XX PN DE19540515-C1.

XX PD 06-FEB-1997.

XX PF 31-OCT-1995; 95DE-1040515.

XX PR 31-OCT-1995; 95DE-1040515.

XX PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX (GESL ) FORSCHUNGSZENTRUM KARLSRUHE GMBH.

XX PI Hekele A, Herrlich P, Ponta H;

XX DR WPI; 1997-101136/10.

XX PT T cells expressing fusion protein specific for variant CD44 gene -  
 XX useful for cancer therapy, esp. pancreatic carcinoma

XX PS Claim 6; Page 9; 16pp; English.

XX CC The fusion protein (T62573) comprises a first portion having  
 CC specific affinity for an amino acid sequence encoded by a variant  
 CC exon of the CD44 gene (pref. W01827 [rat] or W01828 [human]), and a  
 CC second portion comprising at least part of the amino acid sequence of  
 CC a subunit of the T-cell receptor complex or of an immunoglobulin  
 CC receptor. Products contg. the fusion protein (or DNA encoding it)  
 CC are useful for treating cancer, esp. mammary, colonic, gastric or  
 CC pancreatic carcinoma, and metastatic disease.

XX SQ Sequence 14 AA;

Query Match 31.4%; Score 27; DB 18; Length 14;  
 Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENORQGAEP 9

Db 4 enewqgknp 12

Search completed: January 17, 2001, 13:36:43  
 Job time: 118 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:38:36 ; Search time 20.45 Seconds  
(without alignments)  
53.125 Million cell updates/sec

Title: US-08-765-837-5

Perfect score: 86

Sequence: 1 ENQRQGAEPKRNFKVP 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 2778

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_66: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	27.9	15	I38032	hypothetical MNL/T
2	22	25.6	11	D37196	bradykinin-potent
3	22	25.6	11	I49407	placental calcium-
4	22	25.6	15	PA0046	protein QA100044 -
5	22	25.6	15	S67975	apolipoprotein Cb2
6	22	25.6	15	S03955	acidic fibroblast
7	22	25.6	16	A59042	alpha-conotoxin Ep
8	21	24.4	11	XASNBA	bradykinin-potent
9	21	24.4	11	S42587	cell protein - Esc
10	21	24.4	14	C40944	hypothetical prote
11	20	23.3	7	P00777	NADH dehydrogenase
12	20	23.3	11	S68637	acetylcholinestera
13	20	23.3	11	PD0441	translation elonga
14	20	23.3	12	A53252	pollen major aller
15	20	23.3	13	XAV19B	angiotensin-conver
16	20	23.3	13	A37196	bradykinin-potent
17	20	23.3	13	G22565	R-phycocerythrin ga
18	20	23.3	13	PS0443	potassium channel
19	20	23.3	15	SFRF	scotophobin - rat
20	20	23.3	15	A35417	28K serine protein
21	20	23.3	15	S29174	D-galactose-bindin
22	20	23.3	16	A28144	ribosomal protein
23	19	22.1	9	S35538	ribosomal protein
24	19	22.1	9	A41978	calliferramide 1 -
25	19	22.1	9	A44787	calliferramide 10
26	19	22.1	9	D41978	calliferramide 4 -
27	19	22.1	12	A43975	locustamyotropin -
28	19	22.1	12	S65409	histone H2B - huma
29	19	22.1	12	B56049	urinary tract ston

#### ALIGNMENTS

RESULT 1

I38032

hypothetical MNL/TEL mutant fusion protein type I - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 20-Apr-2000

C:Accession: I38032

R:Buijs, A.; Sherr, S.; van Baal, S.; van Bezouw, S.; van der Plas, D.; Van Kessel, A

Oncogene 10, 1511-1519, 1995

A:Title: Translocation (12;22) (p13;q11) in myeloproliferative disorders results in f

A:Reference number: I38031; MUID:95249265

A:Accession: I38032

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-15 <BUI>

A:Cross-references: EMBL:X85025; NID:g971467; PIDN:CAA59398.1; PID:g971468

C:Comment: This sequence is the chimeric product of a translocation mutation.

C:Genetics:

A:Gene: MNL/ETV6; MNL/TEL

A:Map position: 22q11/12p13

C:Keywords: fusion protein

Query Match 27.9%; Score 24; DB 4; Length 15;

Best Local Similarity 37.5%; Pred. No. 6.9e+02;

Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 NQRQGAEP 9

DB 4 NSKEGLQP 11

RESULT 2

D37196

bradykinin-potentiating peptide 4 - island jararaca

C:Species: Bothrops insularis (island jararaca)

C:Date: 14-Feb-1992 #sequence\_revision 01-Dec-1992 #text\_change 05-Aug-14; 1

C:Accession: D37196

R:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.

J. Protein Chem. 9, 221-227, 1990

A:Title: Primary structure and biological activity of bradykinin potentiating peptide

A:Reference number: A37196; MUID:90351557

A:Accession: D37196

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-11 <CIN>

C:Keywords: pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match

Best Local Similarity 25.6%; Score 22; DB 2; Length 11;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 QGAEPR 10  
|| ||  
Db 1 QGGPPR 6

## RESULT 3

I49407  
placental calcium-binding protein - western wild mouse (fragment)  
C:Species: Mus spretus (western wild mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 29-Sep-1999  
C:Accession: I49407  
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.  
Mamm. Genome 5, 349-355, 1994  
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
A:Reference number: I48934; MUID:94319082  
A:Accession: I49407  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-15 <RES>  
A:Cross-references: EMBL:U05696; NID:g497016; PID:AAA61936.1; PID:g497017  
C:Superfamily: S-100 protein; calmodulin repeat homology  
C:Keywords: calcium binding; EF hand

Query Match 25.6%; Score 22; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EPRK 11  
||||  
Db 11 EPRK 14

## RESULT 4

PA0046  
protein QA100044 - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 06-Jun-1997  
C:Accession: PA0046; PA0042  
R:Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
Submitted to JIPID, July 1994  
A:Description: Separation and characterization of Arabidopsis proteins by two-dimensional  
A:Reference number: PA0001  
A:Accession: PA0046  
A:Molecule type: protein  
A:Residues: 1-15 <KAM>  
A:Experimental source: stem

Query Match 25.6%; Score 22; DB 2; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 AEPRKNFV 14  
||:|:  
Db 1 AESKKGFL 8

## RESULT 5

S67975  
apolipoprotein Cb2 - goose (fragment)  
C:Species: Anser anser (domestic goose)  
C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997  
C:Accession: S67975  
R:Hermier, D.; Sellier, N.; Rousselot-Pailley, D.; Forgez, P.  
Eur. J. Biochem. 234, 586-591, 1995  
A:Title: Characterization of apolipoproteins B-100, AI and C from plasma lipoprotein in  
A:Reference number: S67972; MUID:96128192  
A:Accession: S67975  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <HER>

Query Match 25.6%; Score 22; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAEP 9  
||||  
Db 7 GAEP 10

## RESULT 6

S03955  
acidic fibroblast growth factor - dog (fragment)  
N:Alternate names: alpha-endothelial cell growth factor  
C:Species: Canis lupus familiaris (dog)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 18-Jun-1993  
C:Accession: S03955  
R:Quinkler, W.; Maasberg, M.; Bernotat-Danielowski, S.; Luethe, N.; Sharma, H.S.; Sch  
Eur. J. Biochem. 181, 67-73, 1989  
A:Title: Isolation of heparin-binding growth factors from bovine, porcine and canine  
A:Reference number: S03953; MUID:89231704  
A:Accession: S03955  
A:Molecule type: protein  
A:Residues: 1-15 <QUI>  
C:Keywords: growth factor

Query Match 25.6%; Score 22; DB 2; Length 15;  
Best Local Similarity 60.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 NFVKP 16  
|::||  
Db 1 NYMKP 5

## RESULT 7

A59042  
alpha-conotoxin Epi - cone shell (Conus episcopatus)  
C:Species: Conus episcopatus (bishop's cone)  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 13-Aug-1999  
C:Accession: A59042  
R:Loughnan, M.; Bond, T.; Atkins, A.; Cuevas, J.; Adams, D.J.; Broxton, N.M.; Livett,  
J. Biol. Chem. 273, 15667-15674, 1998  
A:Title: Alpha-conotoxin Epi, a novel sulfated peptide from Conus episcopatus that se  
A:Reference number: A59042; MUID:96288307  
A:Accession: A59042  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-16 <LOU>  
C:Superfamily: alpha-conotoxin  
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neu  
F:1-16/Product: alpha-conotoxin Epi #status experimental <MAT>  
F:2-8,3-16/Disulfide bonds: #status experimental  
F:15/Binding site: sulfate (Tyr) (covalent) #status experimental  
F:16/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 25.6%; Score 22; DB 2; Length 16;  
Best Local Similarity 40.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 AEPRKNFVKP 16  
::||| |  
Db 4 SDPRCNMNP 13

## RESULT 8

XASNBA  
bradykinin-potentiating peptide B - mamushi  
C:Species: Agkistrodon blomhoffi (mamushi)  
C:Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 08-Dec-1995

C:Accession: A01254  
 R:Kato, H.; Suzuki, T.  
 Proc. Jpn. Acad. 46, 176-181, 1970  
 A:Reference number: A01254  
 A:Accession: A01254  
 A:Molecule type: protein  
 A:Residues: 1-11 <KAT>  
 A:Note: the sequence of the natural peptide was confirmed by the synthesis and analysis  
 C:Superfamily: bradykinin-potentiating peptide  
 C:Keywords: angiotensin-converting enzyme inhibitor; bradykinin; pyrroglutamic acid; vena  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 24.4%; Score 21; DB 1; Length 11;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 QGAEP R 10  
 II II  
 DB 1 QGLPP R 6

## RESULT 9

S42587  
 celf protein - Escherichia coli (fragment)  
 C:Species: Escherichia coli  
 C:Date: 25-Dec-1994 #sequence\_revision 10-Nov-1995 #text\_change 10-Nov-1995  
 C:Accession: S42587

R:Guzzo, A.; DuBow, M.S.  
 Mol. Gen. Genet. 242, 455-460, 1994  
 A:Title: A luxAB transcriptional fusion to the cryptic celf gene of Escherichia coli dis  
 A:Reference number: S42587; MUID:94166755  
 A:Accession: S42587  
 A:Molecule type: DNA  
 A:Residues: 1-11 <Guz>  
 C:Genetics:  
 A:Gene: celf

Query Match 24.4%; Score 21; DB 2; Length 11;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 ROGAEP 9  
 I I I I  
 DB 2 RRGATP 7

## RESULT 10

C40944  
 hypothetical protein (1-phosphofructokinase 5' region) - Xanthomonas campestris pv. campe  
 C:Species: Xanthomonas campestris pv. campestris  
 C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 18-Jun-1999  
 C:Accession: C40944  
 R:de Crecy-Lagard, V.; Bouvet, O.M.M.; Lejeune, P.; Danchin, A.  
 J. Biol. Chem. 266, 18154-18161, 1991  
 A:Title: Fructose catabolism in Xanthomonas campestris pv. campestris. Sequence of the F  
 A:Reference number: A40944; MUID:92011547  
 A:Accession: C40944  
 A:Molecule type: DNA  
 A:Residues: 1-14 <DE3>

A:Cross-references: GB:M69242; NID:gl55366; PIDN:AAA27600.1; PID:gl55367  
 C:Superfamily: fructose phosphotransferase multiphosphoryltransfer protein; phosphotrans  
 sphotransferase system phosphohistidine-containing protein homology

Query Match 24.4%; Score 21; DB 2; Length 14;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQROG 6  
 I I I I  
 DB 7 EAQREG 12

## RESULT 11

PQ0777  
 NADH dehydrogenase (EC 1.6.99.3) 49K chain - fava bean mitochondrion (fragment)  
 N:Alternate names: complex I 49K chain; NADH--ubiquinone reductase 49K chain  
 C:Species: mitochondrion Vicia faba (fava bean)  
 C:Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 17-Mar-1999  
 C:Accession: PQ0777

R:Letierme, S.; Boutry, M.  
 Plant Physiol. 102, 435-443, 1993  
 A:Title: Purification and preliminary characterization of mitochondrial complex I (NA  
 A:Reference number: PQ0775; MUID:94151437  
 A:Accession: PQ0777

A:Molecule type: protein  
 A:Residues: 1-7 <LET>  
 C:Comment: Complex I, mitochondrial NADH-ubiquinone reductase, is the first of the  
 ranging from 5K to 75K.  
 C:Comment: This enzyme catalyzes electron transfer from endogenous NADH to ubiquinone

C:Genetics:  
 A:Genome: mitochondrion  
 C:Keywords: electron transfer; mitochondrion; oxidoreductase

Query Match 23.3%; Score 20; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ROGA 7  
 I I I I  
 DB 2 ROGA 5

## RESULT 12

S68637  
 acetylcholinesterase (EC 3.1.1.7) P chain - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 04-Dec-1997 #sequence\_revision 04-Dec-1997 #text\_change 30-Jan-1998  
 C:Accession: S68637

R:Boschetti, N.; Brodbeck, U.  
 FEBS Lett. 380, 133-136, 1996  
 A:Title: The membrane anchor of mammalian brain acetylcholinesterase consists of a s  
 A:Reference number: S68637; MUID:96181683  
 A:Accession: S68637

A:Molecule type: protein  
 A:Residues: 1-11 <BOS>  
 A:Experimental source: brain  
 C:Keywords: carboxylic ester hydrolase; glycoprotein; membrane protein

Query Match 23.3%; Score 20; DB 2; Length 11;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 EPRKNEVK 15  
 I I I I I  
 DB 1 EPOKSCSK 8

## RESULT 13

PD0441  
 translation elongation factor TU-like protein P43, mitochondrial - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 21-Aug-1998  
 C:Accession: PD0441

R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.  
 submitted to JIPID, August 1998  
 A:Description: Proteome analysis of mouse brain.  
 A:Reference number: PD0441  
 A:Accession: PD0441

A:Molecule type: protein  
 A:Residues: 1-11 <KAW>  
 A:Experimental source: striatum

C:Keywords: mitochondrion

Query Match 23.3%; Score 20; DB 2; Length 11;  
Best Local Similarity 37.5%; Pred. No. 2.4e+03;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 EPRKNEVK 15  
I : I :  
Db 3 EAKKTYVR 10

# RESULT 14

A53252  
pollen major allergen Par o 1 - Parietaria officinalis (fragment)  
C:Species: Parietaria officinalis  
C:Date: 02-May-1994 #sequence\_revision 25-Apr-1997 #text\_change 25-Apr-1997  
C:Accession: A53252  
R:Oreste, U.; Coscia, M.R.; Scotto d'Abusco, A.; Santonastaso, V.; Ruffilli, A.  
Int. Arch. Allergy Appl. Immunol. 96, 19-27, 1991  
A:Title: Purification and characterization of Par o 1, major allergen of Parietaria officinalis  
A:Reference number: A53252; MUID:92091083  
A:Accession: A53252  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-12 <ORE>  
A:Note: sequence extracted from NCBI backbone (NCBIP:72155)  
A:Note: 5-Thr, 8-Asp, and 9-Ile were also found

Query Match 23.3%; Score 20; DB 2; Length 12;  
Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QGAEP 9  
I I I  
Db 7 QGAMP 11

# RESULT 15

XAVI9B  
angiotensin-converting enzyme inhibitor V-9 - jararaca  
C:Species: Bothrops jararaca (jararaca)  
C:Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 08-Dec-1995  
C:Accession: A01253  
R:Ondetti, M.A.; Williams, N.J.; Sabo, E.F.; Pluscec, J.; Weaver, E.R.; Kocy, O.  
Biochemistry 10, 4033-4039, 1971  
A:Title: Angiotensin-converting enzyme inhibitors from the venom of Bothrops jararaca.  
A:Reference number: A90356; MUID:72118526  
A:Accession: A01253  
A:Molecule type: protein  
A:Residues: 1-13 <OND>  
A:Note: the structure of the peptide was confirmed by synthesis  
C:Comment: This peptide also potentiates bradykinin by inhibiting the kinases that inactivate bradykinin.  
C:Superfamily: bradykinin-potentiating peptide  
C:Keywords: angiotensin-converting enzyme inhibitor; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 23.3%; Score 20; DB 1; Length 13;  
Best Local Similarity 66.7%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 QGAEP 10  
I I I  
Db 1 QGGWPR 6

Search completed: January 17, 2001, 13:38:37  
Job time: 167 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:39:28 ; Search time 12.32 Seconds  
(without alignments)  
41.940 Million cell updates/sec

Title: US-08-765-837-5  
Perfect score: 86  
Sequence: 1 ENORQGAERKRFVKP 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues  
Total number of hits satisfying chosen parameters: 777

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	26	30.2	8	1 UC26_MAIZE	P80632 zea mays (m
2	23	26.7	13	1 BP37_LEUMA	P81754 leucophaea
3	23	26.7	14	1 GLGS_SPIOL	P55235 spinacia ol
4	23	26.7	15	1 UC28_MAIZE	P80634 zea mays (m
5	22	25.6	11	1 BP4_BOTIN	P30424 bothrops in
6	22	25.6	15	1 FGFI_CANFA	P18651 canis faml
7	22	25.6	16	1 CXAI_CONEP	P56638 conus episc
8	21	24.4	11	1 BPPB_AGKHA	P01021 agkistrodon
9	21	24.4	14	1 YMPX_XANCP	Q03397 xanthomonas
10	20	23.3	13	1 BP1_BOTJA	P01020 bothrops ja
11	20	23.3	13	1 LPAA_PORGI	P81411 porphyromon
12	20	23.3	15	1 LEC3_AXIPO	P28588 axinella po
13	20	23.3	15	1 SCOT_RAT	P01159 rattus norv
14	19	22.1	8	1 RS7_MYCIT	P33564 mycobacteri
15	19	22.1	9	1 FAR1_CALVO	P41856 calliphora
16	19	22.1	9	1 FAR4_CALVO	P41856 calliphora
17	19	22.1	9	1 FARA_CALVO	P41865 calliphora
18	19	22.1	12	1 LMT1_LOEMI	P22395 locusta mig
19	19	22.1	15	1 CHI1_PEA	P21225 pisum sativ
20	18	20.9	7	1 FAR4_PANRE	P41875 panagrellus
21	18	20.9	8	1 RS10_SALTY	O68928 salmonella
22	18	20.9	9	1 RS10_SERMA	O68936 serratia ma
23	18	20.9	10	1 AMPN_HELAM	P81731 helicoverpa
24	18	20.9	11	1 TKN_PHYFU	P08615 physalaemus
25	18	20.9	12	1 FAR1_CALVO	P41869 calliphora
26	18	20.9	13	1 MPL_MICOC	P81532 microplitis
27	18	20.9	15	1 IRBP_CRISP	P12665 cricetidae
28	18	20.9	15	1 UC16_MAIZE	P80622 zea mays (m
29	17	19.8	7	1 FAR1_PROCL	P38499 procambarus
30	17	19.8	7	1 FAR2_PROCL	P38498 procambarus
31	17	19.8	8	1 ALL1_CVDPO	P82152 cydia pomon
32	17	19.8	8	1 FAR3_HOMAM	P41486 homarus ame
33	17	19.8	8	1 FAR4_HOMAM	P41487 homarus ame

34 17 19.8 8 1 UH09\_RAT P56575 rattus norv  
35 17 19.8 9 1 FAR2\_CALVO P41857 calliphora  
36 17 19.8 9 1 FAR2\_PANRE P41873 panagrellus  
37 17 19.8 9 1 FAR3\_CALVO P41858 calliphora  
38 17 19.8 11 1 LADD\_ONCMY P81018 oncorhynchu  
39 17 19.8 13 1 BOML\_PSEGU P42991 pseudophryn  
40 17 19.8 13 1 CXAA\_CONST P28878 conus stria  
41 17 19.8 13 1 FARB\_ASCSU P43173 ascaris suu  
42 17 19.8 13 1 UPS1\_UPEIN P82036 uperoleia l  
43 17 19.8 14 1 EFTU\_CANFA P54835 canis faml  
44 17 19.8 14 1 RS19\_PFWBP O52093 pigeon pea  
45 17 19.8 14 1 SMS1\_MYOSC P20750 myoxocsphal

## ALIGNMENTS

RESULT 1  
UC26\_MAIZE  
ID UC26\_MAIZE STANDARD; PRT; 8 AA.  
AC P80632;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 907).  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=COLEOPTILE;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
RA Pernollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
RT genome analysis program";  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 7.0, ITS MW IS: 57.2 KDA.  
DR MAIZE-2DPAGE; P80632; COLEOPTILE.  
FT NON\_TER 1  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 990 MW; 9639D6DAB4176B1D CRC64;

Query Match 30.2%; Score 26; DB 1; Length 8;  
Best Local Similarity 71.4%; Pred. No. 8.9e+04;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AEPRKNF 13  
Db 1 AEPRDQF 7

RESULT 2  
BP37\_LEUMA  
ID BP37\_LEUMA STANDARD; PRT; 13 AA.  
AC P81754;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE BRAIN PROTEIN 37F3.  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blaberidae; Blaberidae; Leucophaea.  
RN [1]  
RP SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE=BRAIN;  
RX MEDLINE=97269266; PubMed=9114447;  
RA Muren J.E., Naessel D.R.;  
RT "Seven tachykinin-related peptides isolated from the brain of the  
RT madeira cockroach; evidence for tissue-specific expression of

RT isoforms.";  
 RL Peptides 18:7-15(1997).  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- MASS SPECTROMETRY: MW=1435.7; METHOD=MALDI-MS.  
 SQ SEQUENCE 13 AA; 1436 MW; 8E532C9DE9A7D1B5 CRC64;

Query Match 26.7%; Score 23; DB 1; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 3.4e+02;  
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENQROGAEP 10  
 | | | | |  
 Db 4 ESTVSAEP 13

RESULT 3  
 GLGS\_SPIOL STANDARD; PRT; 14 AA.  
 ID GUGS\_SPIOL  
 AC P5235;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE GLUCOSE-1-PHOSPHATE ADENYLTRANSFERASE SMALL SUBUNIT (EC 2.7.7.27)  
 DE (ADP-GLUCOSE SYNTHASE) (ADP-GLUCOSE PYROPHOSPHORYLASE) (AGPASE B)  
 DE (ALPHA-D-GLUCOSE-1-PHOSPHATE ADENYL TRANSFERASE) (FRAGMENT).  
 OS Spinacia oleracea (Spinach).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;  
 CC Caryophyllales; Chenopodiaceae; Spinacia.  
 RN [1]  
 RP SEQUENCE.

RC TISSUE=LEAF;  
 RA Morell M.K., Bloom M., Knowles V., Preiss J.;  
 RT "Subunit structure of spinach leaf ADPglucose pyrophosphorylase.";  
 RL Plant Physiol. 85:182-187(1987).  
 CC -1- FUNCTION: THIS PROTEIN PLAYS A ROLE IN SYNTHESIS OF STARCH.  
 CC IT CATALYSES THE SYNTHESIS OF THE ACTIVATED GLYCOSYL DONOR,  
 CC ADP-GLUCOSE FROM GLC-1-P AND ATP.  
 CC -1- CATALYTIC ACTIVITY: ATP + ALPHA-D-GLUCOSE 1-PHOSPHATE =  
 CC PYROPHOSPHATE + ADP-GLUCOSE.  
 CC -1- ENZYME REGULATION: ACTIVATED BY 3'PHOSPHOGLYCERATE, INHIBITED  
 CC BY ORTHOPHOSPHATE. ALLOSTERIC REGULATION.  
 CC -1- PATHWAY: STARCH BIOSYNTHESIS.  
 CC -1- SUBUNIT: HETEROTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLASTS OF LEAVES AND AMYLOPLASTS  
 CC OF DEVELOPING ENDOSPERM.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-  
 CC PHOSPHATE ADENYLYLTRANSFERASE FAMILY.

DR INTERPRO: IPR001825; -;  
 DR PROSITE: PS00808; ADP\_GLC\_PYROPHOSPH\_1; PARTIAL.  
 DR PROSITE: PS00809; ADP\_GLC\_PYROPHOSPH\_2; PARTIAL.  
 DR PROSITE: PS00810; ADP\_GLC\_PYROPHOSPH\_3; PARTIAL.  
 KW Glycogen biosynthesis; Transferase; Nucleotidyltransferase;  
 KW Multigene family; Starch biosynthesis; Allosteric enzyme;  
 KW Amyloplast; Chloroplast.  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1490 MW; 98B5792C3AE738C5 CRC64;

Query Match 26.7%; Score 23; DB 1; Length 14;  
 Best Local Similarity 33.3%; Pred. No. 3.6e+02;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENQROGAEP 9  
 : | : | : |  
 Db 5 QNSODGLDP 13

RESULT 4  
 UC28\_MAIZE STANDARD; PRT; 15 AA.  
 ID UC28\_MAIZE  
 AC P80634;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 984)  
 DE (FRAGMENT).  
 OS Zea mays (Maize).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=COLEOPTILE;  
 RA touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RA Pernollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize two dimensional gel protein database: towards an integrated  
 RT genome analysis program.";  
 RL Theor. Appl. Genet. 93:997-1005(1996).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.2, ITS MW IS: 27.7 KDA.  
 CC MAIZE-2DPAGE; P80634; COLEOPTILE.  
 DR MAIZEDB; 123959; -;  
 DR MAIZEDB; 123959; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1631 MW; 45C554E40BE9E77F CRC64;

Query Match 26.7%; Score 23; DB 1; Length 15;  
 Best Local Similarity 71.4%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NQROGAE 8  
 | | | | |  
 Db 9 NQSDAE 15

RESULT 5  
 BPP4\_BOTIN STANDARD; PRT; 11 AA.  
 ID BPP4\_BOTIN  
 AC P30424;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE BRADYKININ-POTENTIATING PEPTIDE S4,1,2 (ANGIOTENSIN-CONVERTING  
 DE ENZYME INHIBITOR).  
 OS Bothrops insularis (Island Jararaca) (Queimada Jararaca).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 CC Viperidae; Crotalinae; Bothrops.  
 RN [1]  
 RP SEQUENCE.

RC TISSUE=VENOM;  
 RX MEDLINE=90351557; PubMed=2386615;  
 RA Cintra A.C.O., Vieira C.A., Giglio J.R.;  
 RT "Primary structure and biological activity of bradykinin potentiating  
 RT peptides from Bothrops insularis snake venom.";  
 RL J. Protein Chem. 9:221-227(1990).  
 CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE  
 CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF  
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.  
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.

DR PIR; D37196; D37196.  
 KW Hypotensive agent; Venom.  
 FT MOD\_RES 1 1  
 SQ SEQUENCE 11 AA; 1143 MW; 20BBBF13C7741777 CRC64;

Query Match 25.6%; Score 22; DB 1; Length 11;  
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 QGAEP 10  
 | | | | |  
 Db 1 QGGP 6

RESULT 6  
FGF1\_CANFA  
ID FGF1\_CANFA STANDARD; PRT; 15 AA.  
AC P18651;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE HEPARIN-BINDING GROWTH FACTOR 1 (HBGF-1) (ACIDIC FIBROBLAST GROWTH FACTOR) (AFGF) (ALPHA-ENDOTHELIAL CELL GROWTH FACTOR) (FRAGMENT).  
GN FGF1 OR FGF-1.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
RN [1]  
RP SEQUENCE  
RX MEDLINE=89231704; PubMed=2714282;  
RA Quinkler W., Maasberg M., Bernotat-Danielowski S., Luethke N., Sharma H.S., Schaper W.;  
RT "Isolation of heparin-binding growth factors from bovine, porcine and canine hearts";  
RL Eur. J. Biochem. 181:67-73(1989).  
CC -!- FUNCTION: THE HEPARIN-BINDING GROWTH FACTORS ARE ANGIOGENIC AGENTS IN VIVO AND ARE POTENT MITOGENS FOR A VARIETY OF CELL TYPES IN VITRO. THERE ARE DIFFERENCES IN THE TISSUE DISTRIBUTION AND CONCENTRATION OF THESE 2 GROWTH FACTORS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS HEPARIN, ALTHOUGH LESS STRONGLY THAN DOES HBGF.  
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.  
DR HIR; S03955; S03955.  
DR HSP; P05230; 2AXM.  
DR INTERPRO; IPR002209;  
DR PROSITE; PS00247; HBGF\_FGF; PARTIAL.  
KW Growth factor; Mitogen; Vascularization; Heparin-binding.  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1732 MW; 53CC9A3CADDDDAA1 CRC64;

Query Match 25.6%; Score 22; DB 1; Length 15;  
Best Local Similarity 60.0%; Pred. No. 5.8e+02;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 NFVKP 16  
DB 1 NYMKP 5

RESULT 7  
CXAL\_CONEP  
ID CXAL\_CONEP STANDARD; PRT; 16 AA.  
AC P56638;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE ALPHA-CONOTOXIN EPI.  
OS Conus episcopatus (Bishop's cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda; Neogastropoda; Conoidea; Conidae; Conus.  
RN [1]  
RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).  
RX MEDLINE=98376423; PubMed=9708977;  
RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H., Alewood P.F., Lewis R.J., Martin J.L.;  
RT "The 1.1-A resolution crystal structure of [Tyr15]Epi, a novel alpha-conotoxin from Conus episcopatus, solved by direct methods.";  
RL Biochemistry 37:11425-11433(1998).  
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-3/BETA-4 SUBUNITS.  
CC PDB; 1A0M; 13-JAN-99.

KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
KW Venom; Sulfatation; 3D-structure.  
FT DISULFID 2 8  
FT DISULFID 3 16  
FT DISULFID 15 15  
FT MOD\_RES 16 16  
SQ SEQUENCE 16 AA; 1792 MW; C63385F376C99B4C CRC64;

Query Match 25.6%; Score 22; DB 1; Length 16;  
Best Local Similarity 40.0%; Pred. No. 6.2e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ABPRKNFVKP 16  
DB 4 SDPRCNMNP 13

RESULT 8  
BPPB\_AKHA  
ID BPPB\_AKHA STANDARD; PRT; 11 AA.  
AC P10211;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE BRADYKININ-POTENTIATING PEPTIDE B (ANGIOTENSIN-CONVERTING ENZYME INHIBITOR).  
OS Agkistrodon halys blomhoffi (Mamushi) (Gloydius blomhoffii).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Agkistrodon.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=VENOM;  
RA Kato H., Suzuki T.;  
RT "Amino acid sequence of bradykinin-potentiating peptide isolated from the venom of Agkistrodon halys blomhoffii";  
RL Proc. Jpn. Acad. B, Phys. Biol. Sci. 46:176-181(1970);  
CC -!- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.  
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.  
DR PIR; A01254; XASNBA.  
KW Hypotensive agent; Venom.  
FT MOD\_RES 1 1  
SQ SEQUENCE 11 AA; 1199 MW; 295CBF0627741777 CRC64;

Query Match 24.4%; Score 21; DB 1; Length 11;  
Best Local Similarity 66.7%; Pred. No. 6.3e+02;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 QCAEPR 10  
DB 1 QGLPPR 6

RESULT 9  
YMTF\_XANCP  
ID YMTF\_XANCP STANDARD; PRT; 14 AA.  
AC Q03397;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-OCT-1993 (Rel. 27, Last annotation update)  
DE HYPOTHETICAL PROTEIN IN FRUK 5' REGION (FRAGMENT).  
GN MTP.  
OS Xanthomonas campestris (pv. campestris).  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; Xanthomonas.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92011547; PubMed=1655739;  
RA de Crecy-Lagard V., Bouvet O.M., Lejeune P., Danchin A.;

RT "Fructose catabolism in *Xanthomonas campestris* pv. *campestris*.  
RT Sequence of the PTS operon, characterization of the fructose-specific  
RL J. Biol. Chem. 266:18154-18161(1991).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M69242; AAA27600.1; -  
DR PIR; C40944; C40944.  
KW Hypothetical protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 14 AA; 1585 MW; 72EA9AA3CE16CCDB CRC64;  
  
Query Match 24.4%; Score 21; DB 1; Length 14;  
Best Local Similarity 66.7%; Pred. No. 8.1e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ENORQG 6  
| | | |  
DB 7 EAQREG 12  
  
RESULT 10  
BPPI\_BOTJA STANDARD; PRT; 13 AA.  
AC P01020; P30421;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE BRADYKININ-POTENTIATING PEPTIDE S3.1 (13A) (ANGIOTENSIN-CONVERTING  
DE ENZYME INHIBITOR V-9).  
OS Bothrops jararaca (Jararaca), and  
OS Bothrops insularis (Island jararaca) (Queimada jararaca).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;  
OC Viperidae; Crotalinae; Bothrops.  
RN [1]  
RP SEQUENCE.  
RC SPECIES=B. JARARACA; TISSUE=VENOM;  
RX MEDLINE=72118526; PubMed=4334402;  
RA Ondetti M.A., Williams N.J., Sabo E.F., Pluscec J., Weaver E.R.,  
RA Kocy O.;  
RT "Angiotensin-converting enzyme inhibitors from the venom of Bothrops  
RT jararaca. Isolation, elucidation of structure, and synthesis.";  
RL Biochemistry 10:4033-4039(1971).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=B. INSULARIS; TISSUE=VENOM;  
RX MEDLINE=90351557; PubMed=2386615;  
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;  
RT "Primary structure and biological activity of bradykinin potentiating  
RT peptides from Bothrops insularis snake venom.";  
RL J. Protein Chem. 9:221-227(1990).  
CC -!- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE  
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF  
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.  
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.  
DR PIR; A01253; XAV19B.  
DR PIR; A37196; A37196.  
KW Hypotensive agent; Venom.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
SQ SEQUENCE 13 AA; 1388 MW; 6824FC97D83D6774 CRC64;  
  
Query Match 23.3%; Score 20; DB 1; Length 13;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 QGAEPK 10  
| | | |  
DB 1 QGGWPR 6  
  
RESULT 11  
LPAA\_PORGI STANDARD; PRT; 13 AA.  
AC P81411;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE LIPID-A-ASSOCIATED PROTEIN (FRAGMENT).  
OS Porphyromonas gingivalis (Bacteroides gingivalis).  
OC Bacteria; CFB group; Bacteroidaceae; Porphyromonadaceae.  
RN [1]  
RP SEQUENCE.  
RC STRAIN=W50;  
RX MEDLINE=99061194; PubMed=9846737;  
RA Sharp L., Poole S., Reddi K., Fletcher J., Nair S., Wilson M.,  
RA Curtis M., Henderson B., Tabona P.;  
RT "A lipid A-associated protein of Porphyromonas gingivalis, derived  
RT from the haemagglutinating domain of the RI protease gene family, is  
RT a potent stimulator of interleukin 6 synthesis.";  
RL Microbiology 144:3019-3026(1998).  
CC -!- FUNCTION: IS ASSOCIATED WITH LIPID A, A PHOSPHORYLATED GLYCOLIPID  
CC THAT ANCHORS THE LIPOPOLYSACCHARIDE TO THE OUTER MEMBRANE OF THE  
CC CELL.  
CC VARIANT 12 12 G -> F.  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1346 MW; 38EA796EAF63AB7 CRC64;  
  
Query Match 23.3%; Score 20; DB 1; Length 13;  
Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 5 QGAEPK 12  
| | | |  
DB 2 QGDNPKD 9  
  
RESULT 12  
LEC3\_AXIPO STANDARD; PRT; 15 AA.  
AC P28588;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-DEC-1992 (Rel. 24, Last annotation update)  
DE LECTIN III (FRAGMENT).  
OS Axinella polypoides (Sponge).  
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;  
OC Halichondrida; Axinellidae; Axinella.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93003351; PubMed=1390906;  
RA Buck F., Luth C., Strupat K., Bretting H.;  
RT "Comparative investigations on the amino-acid sequences of different  
RT isolectins from the sponge Axinella polypoides (Schmidt).";  
RL Biochim. Biophys. Acta 1159:1-8(1992).  
CC -!- FUNCTION: POSSESSES D-GALACTOSE BINDING SPECIFICITY.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SUBCELLULAR LOCATION: STORED IN SPHERULOUS CELLS IN THE SPONGE  
CC TISSUE.  
KW Lectin.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1744 MW; DC05E47E3F8DC5F3 CRC64;  
  
Query Match 23.3%; Score 20; DB 1; Length 15;  
Best Local Similarity 42.9%; Pred. No. 1.3e+03;

Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 AEPRKNF 13  
| | | | |  
Db 1 ASPXENY 7

RESULT 13  
SCOT\_RAT STANDARD; PRT; 15 AA.  
AC P01159;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE SCOTOPHOBIN.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RX MEDLINE=72240741; PubMed=4558348;  
RA Ungar G., Desiderio D.M., Parr W.;  
RT "Isolation, identification and synthesis of a specific-behaviour  
inducing brain peptide.";  
RL Nature 238:198-202(1972).  
RN [2]  
RP SEQUENCE.  
RA Desiderio D.M., Ungar G., White P.A.;  
RN J. Chem. Soc. Chem. Commun. 432-433(1971).  
RN [3]  
RP REFERENCE'S COMMENTS ON REF.1.  
RX MEDLINE=72240742; PubMed=4558349;  
RA Stewart W.W.;  
RT "Comments on the chemistry of scotophobin.";  
RL Nature 238:202-209(1972).  
RN [4]  
RP REPLY TO REFERENCE'S COMMENTS.  
RA Ungar G., Desiderio D.M., Parr W.;  
RL Nature 238:209-210(1972).  
CC -1- TISSUE SPECIFICITY: BRAIN.  
DR PIR; A01423; SFRT.  
KW Amidation.  
FT MOD.RES 15 15 AMIDATION.  
SQ SEQUENCE 15 AA; 1582 MW; 5B87696F6C7FD76C CRC64;

Query Match 23.3%; Score 20; DB 1; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENOROG 6  
| | | | |  
Db 2 DNNQQG 7

RESULT 14  
RS7\_MYCIT STANDARD; PRT; 8 AA.  
AC P33564;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE 30S RIBOSOMAL PROTEIN S7 (FRAGMENT).  
GN RPSG.  
OS Mycobacterium intracellulare.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93197130; PubMed=8451173;  
RA Nair J., Rouse D.A., Morris S.L.;  
RT "Nucleotide sequence analysis of the ribosomal S12 gene of  
Mycobacterium intracellulare.";

RL Nucleic Acids Res. 21:1039-1039(1993).  
CC -1- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF  
16S RIBOSOMAL RNA (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L08171; AAA25376.1; -  
DR PIR; S35538; S35538.  
DR INTERPRO; IPR000235; -  
DR PROSITE; PS00052; RIBOSOMAL\_S7; PARTIAL.  
KW Ribosomal protein; rRNA-binding.  
FT INIT\_MET 0 BY SIMILARITY.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 850 MW; 63276DC768732417 CRC64;

Query Match 22.1%; Score 19; DB 1; Length 8;  
Best Local Similarity 42.9%; Pred. No. 8.9e+04;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RQAEPR 10  
| | | | |  
Db 2 RKGPAK 8

RESULT 15  
FARL\_CALVO STANDARD; PRT; 9 AA.  
AC P41856;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CALLIFMRFAMIDE 1.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=THORACIC GANGLION;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
RA Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
neuropeptides (designated callifmrfamides) from the blowfly  
Calliphora vomitoria.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED  
SALIVARY GLAND OF CALLIPHORA.  
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
FAMILY.  
CC PIR; A41978; A41978.  
DR Neuropeptide; Amidation.  
FT MOD.RES 9  
SQ SEQUENCE 9 AA; 1169 MW; 29D00699CAB6C6C7 CRC64;

Query Match 22.1%; Score 19; DB 1; Length 9;  
Best Local Similarity 28.6%; Pred. No. 8.9e+04;  
Matches 2; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 9 PRKNFVK 15  
| | | | |  
Db 2 PQQDFMR 8

Search completed: January 17, 2001, 13:39:29  
Job time: 164 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:39:12 ; Search time 32.75 Seconds  
(without alignments)  
57.262 Million cell updates/sec

Title: US-08-765-837-5

Perfect score: 86

Sequence: 1 ENQRQGAERPRKRVKVP 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 2612

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: Sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	27.9	11	5	P81687
2	24	27.9	13	4	Q14804
3	24	27.9	16	6	Q9TQ29
4	23	26.7	16	6	Q9TRB4
5	22	25.6	10	11	Q9QVE6
6	22	25.6	12	12	Q68772
7	22	25.6	13	2	Q9K9D7
8	22	25.6	15	8	P82135
9	22	25.6	15	11	Q62531
10	22	25.6	16	2	P82597
11	22	25.6	16	2	Q9KC50
12	22	25.6	16	4	Q9NNZ2
13	21	24.4	11	11	Q62207
14	21	24.4	12	12	Q9ICE5
15	21	24.4	15	11	Q9QV76
16	21	24.4	15	13	Q9PS10
17	20	23.3	10	12	Q88082
18	20	23.3	13	8	Q9MOK3
19	20	23.3	13	11	Q63940

20	20	23.3	13	12	Q83171
21	20	23.3	14	5	Q9VN49
22	20	23.3	14	12	O10231
23	20	23.3	14	12	O10234
24	20	23.3	14	12	O10226
25	20	23.3	14	12	O10227
26	20	23.3	14	12	O10228
27	20	23.3	14	12	O10229
28	20	23.3	14	12	O10230
29	20	23.3	14	12	O10232
30	20	23.3	14	12	O10233
31	20	23.3	14	12	O10235
32	20	23.3	15	3	Q9URC5
33	20	23.3	15	6	Q9TRN3
34	20	23.3	15	6	Q9TRH1
35	20	23.3	16	5	Q9TWK0
36	19	22.1	9	2	Q53914
37	19	22.1	10	11	Q9QVE7
38	19	22.1	10	11	Q9JLT5
39	19	22.1	10	12	O90348
40	19	22.1	15	2	O54298
41	19	22.1	15	10	P82331
42	19	22.1	16	2	Q9R596
43	19	22.1	16	4	Q07059
44	19	22.1	16	9	O21922
45	18	20.9	9	8	Q9T2L1

#### ALIGNMENTS

RESULT 1

P81687					
ID	P81687	PRELIMINARY:	PRT:	11	AA.
AC	P81687				
DT	01-NOV-1998 (TREMBLrel. 08, Created)				
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)				
DT	01-MAY-1999 (TREMBLrel. 10, Last annotation update)				
DE	ANDROPIN (FRAGMENT).				
GN	ANP.				
OS	Drosophila sechellia (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7238;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=MONTEVAYOR;				
RX	MEDLINE=98393576; PubMed=9725836;				
RA	Ramos-Onsins S., Aguade M.;				
RT	"Molecular evolution of the Cecropin multigene family in Drosophila:				
RT	functional genes vs pseudogenes."				
RL	Genetics 150:157-171(1998).				
CC	!- FUNCTION: MALE-SPECIFIC PEPTIDE WITH MODERATE ACTIVITY AGAINST				
CC	GRAM-POSITIVE BACTERIA (BY SIMILARITY).				
CC	!- SIMILARITY: STRUCTURALLY RELATED TO CECROPINS.				
DR	EMBL; Y16862; CAA76482.1; -				
DR	FLYBASE; FBgn0025847; DsecVAnp.				
KW	Insect immunity; Antibiotic.				
FT	NON_TER 1				
SQ	SEQUENCE 11 AA; 1254 MW; 0B6F902556944B10 CRC64;				

Query Match 27.9%; Score 24; DB 5; Length 11;  
Best Local Similarity 40.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 AEPKRVKVP 16  
1:1:1:1  
Db 1 AKPIENMILP 10

RESULT 2

Q14804  
ID Q14804 PRELIMINARY; PRT; 13 AA.  
AC Q14804;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE CHLORIDE CHANNEL PROTEIN (FRAGMENT).  
GN CIC-1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95050513; PubMed=7961681;  
RA Klocke R., Steinmeyer K., Jentsch T.J., Jockusch H.;  
RT "Role of innervation, excitability, and myogenic factors in the  
RT expression of the muscular chloride channel CIC-1. A study on normal  
RT and myotonic muscle.";  
RL J. Biol. Chem. 269:27635-27639(1994).  
DR EMBL: Z31375; CAA83246.1; -;  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1530 MW; 9DC880B73FCD16C5 CRC64;

Query Match 27.9%; Score 24; DB 4; Length 13;  
Best Local Similarity 50.0%; Pred. No. 9.3e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ENQRQAE 8  
Db :|||  
5 QSORHGE 12

RESULT 3  
ID Q9TQ29 PRELIMINARY; PRT; 16 AA.  
AC Q9TQ29;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE 20 KDa PHOSPHORYLATION-DEPENDENT PROTEIN PHOSPHATASE-1 INHIBITORY  
DE PROTEIN (FRAGMENT).  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96362144; PubMed=8720121;  
RA Eto M., Ohmori T., Suzuki M., Furuya K., Morita F.;  
RT "A novel protein phosphatase-1 inhibitory protein potentiated by  
RT protein kinase C. Isolation from porcine aorta media and  
RT characterization.";  
RL J. Biochem. 118:1104-1107(1995).  
SQ SEQUENCE 16 AA; 1822 MW; 7711966667B09400 CRC64;

Query Match 27.9%; Score 24; DB 6; Length 16;  
Best Local Similarity 57.1%; Pred. No. 1.1e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 PRKNFK 15  
Db :|||  
5 PTENFV 11

RESULT 4  
ID Q9TRB4 PRELIMINARY; PRT; 16 AA.  
AC Q9TRB4;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE ATP-DEPENDENT 20 S PROTEASOME ACTIVATOR (FRAGMENT).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94342244; PubMed=8063704;  
RA DeMartino G.N., Moomaw C.R., Zagnitko O.P., Proske R.J., Chu-Ping M.,  
RA Afendis S.J., Swaffield J.C., Slaughter C.A.;  
RT "PA700, an ATP-dependent activator of the 20 S proteasome, is an  
RT ATPase containing multiple members of a nucleotide-binding protein  
RT family.";  
RL J. Biol. Chem. 269:20878-20884(1994).  
SQ SEQUENCE 16 AA; 1878 MW; F70F74211EE26EDE CRC64;

Query Match 26.7%; Score 23; DB 6; Length 16;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 EPRKNF 13  
Db :|||  
3 EPEPNF 8

RESULT 5  
ID Q9QVE6 PRELIMINARY; PRT; 10 AA.  
AC Q9QVE6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE PROTAMINE MP2 INTERMEDIATE PROTEIN PMP2/26.  
OS Mus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10095;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92174934; PubMed=1541289;  
RA Chauviere M., Martinage A., Debarle M., Sautiere P., Chevaillier P.;  
RT "Molecular characterization of six intermediate proteins in the  
RT processing of mouse protamine p2 precursor.";  
RL Eur. J. Biochem. 204:759-765(1992).  
SQ SEQUENCE 10 AA; 1028 MW; 2B099C75B72866D8 CRC64;

Query Match 25.6%; Score 22; DB 11; Length 10;  
Best Local Similarity 44.4%; Pred. No. 1.6e+03;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 QRQGAEPK 11  
Db :|||  
2 QGGLSPER 10

RESULT 6  
ID Q68772 PRELIMINARY; PRT; 12 AA.  
AC Q68772;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE (FRAGMENT).  
OS Simian hemorrhagic fever virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage.  
OX NCBI\_TaxID=38143;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LVR 42-0/M6941;



RX MEDLINE=95193253; PubMed=7886957;  
RA Zeng L., Godeny E.K., Methven S.L., Brinton M.A.;  
RT "Analysis of simian hemorrhagic fever virus (SHFV) subgenomic RNAs,  
RT Junction sequences, and 5' leader";  
RL Virology 207:543-548(1995).  
DR EMBL: L39091; AAA68933.1; -.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1398 MW; 88911EA56CA4176E CRC64;

Query Match 25.6%; Score 22; DB 12; Length 12;  
Best Local Similarity 66.7%; Pred. No. 1.9e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 PRKNEV 14  
|||  
DB 6 PRSNLV 11

RESULT 7  
Q9K9D7 PRELIMINARY; PRT; 13 AA.  
AC Q9K9D7;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE BH2711 PROTEIN.  
GN BH2711.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RA Takami H., Nakasone K., Takaki Y.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP001516; BAB06430.1; -.  
SQ SEQUENCE 13 AA; 1606 MW; CAE02359F55A25B9 CRC64;

Query Match 25.6%; Score 22; DB 2; Length 13;  
Best Local Similarity 57.1%; Pred. No. 2.1e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 RKNFVKP 16  
|||  
DB 3 QKRFSKP 9

RESULT 8  
P82135 PRELIMINARY; PRT; 15 AA.  
AC P82135;  
DT 01-JUN-2000 (TReMBLrel. 14, Created)  
DT 01-JUN-2000 (TReMBLrel. 14, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S18 ALPHA (FRAGMENT).  
OS Spinacia oleracea (Spinach).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;  
OC Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV. ALVARO; TISSUE=LEAF;  
RA Yamauchi K., von Knoblauch K., Subramanian A.R.;  
RT "Identification of all the proteins in the small subunit of an  
RT Organelle (chloroplast) ribosome";  
RL J. Biol. Chem. 0:0-0(2000).  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.

CC -1- MISCELLANEOUS: S18 ALPHA AND BETA FORMS DIFFER IN P1. S18 BETA  
CC FORM IS THE LEAST BASIC FORM.  
CC -1- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.  
DR INTERPRO: IPR001648; -.  
DR PROSITE: PS00057; RIBOSOMAL\_S18; PARTIAL.  
KW Ribosomal protein; Chloroplast.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1910 MW; C29396F592403B60 CRC64;

Query Match 25.6%; Score 22; DB 8; Length 15;  
Best Local Similarity 30.8%; Pred. No. 2.4e+03;  
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 ENORQGAEPKNF 13  
:::| :|::|  
DB 3 KSKRPFIKSRNF 15

RESULT 9  
Q62531 PRELIMINARY; PRT; 15 AA.  
AC Q62531;  
DT 01-JUN-1998 (TReMBLrel. 06, Created)  
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
DE PLACENTAL CALCIUM-BINDING PROTEIN (18A2) (PEL98) (WTS1 PROTEIN)  
DE (METASTATIC CELL PROTEIN) (FRAGMENT).  
GN MTS1.  
OS Mus spretus (Western wild mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRET/EI;  
RX MEDLINE=94319082; PubMed=8043949;  
RA Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maezaki Y.,  
RA Nadeau J.H.;  
RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR";  
RL Mamm. Genome 5:349-355(1994).  
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, AND MORE  
CC SPECIFICALLY TO S-100/CABP LIKE PROTEINS.  
DR EMBL: U05696; AAA61936.1; -.  
KW Calcium-binding; Placenta.  
FT NON\_TER 1 1  
SQ SEQUENCE 15 AA; 1824 MW; 31EBA50E3987741 CRC64;

Query Match 25.6%; Score 22; DB 11; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EPRK 11  
||||  
DB 11 EPRK 14

RESULT 10  
P82597 PRELIMINARY; PRT; 16 AA.  
AC P82597;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE THERMOSTABLE MONOACYLGLYCEROL LIPASE (MGLP) (24 KDA) (BC 3.1.1.23)  
DE (FRAGMENT).  
OS Bacillus sp.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=129908;  
RN [1]  
RP SEQUENCE.

RC STRAIN=H-257;  
 RX MEDLINE=20198254; PubMed=10731713;  
 RA Inamura S., Kitaoura S.;  
 RT "Purification and characterization of a monoacylglycerol lipase from  
 the moderately thermophilic *Bacillus* sp. H-257.";  
 RL J. Biochem. 127:419-425(2000).  
 CC -!- FUNCTION: HYDROLYZES MONOACYLGLYCEROLS, WITH THE HIGHEST ACTIVITY  
 OCCURRING WITH 1-MONOLAURYLGLYCEROL.  
 CC -!- ENZYME REGULATION: NOT INHIBITED BY CHOLATE, BUT SLIGHTLY  
 INHIBITED BY TRITON X-100 AND DEOXYCHOLATE.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- MISCELLANEOUS: HAS AN ISOELECTRIC POINT OF 4.66. OPTIMUM  
 TEMPERATURE IS 75 DEGREES CELSIUS AND MAXIMUM ACTIVITY AT PH 6-8.  
 CC -!- MISCELLANEOUS: THIS LIPASE IS NOT SECRETED EXTRACELLULARLY AS  
 OTHER BACTERIAL LIPASES.  
 KW Hydrolase.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1787 MW; 900CF59289521D8F CRC64;

Query Match 25.6%; Score 22; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GAEP 9  
 ||||  
 Db 9 GAEP 12

## RESULT 11

ID Q9KC50 PRELIMINARY; PRT; 16 AA.  
 AC Q9KC50;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE BH1724 PROTEIN.  
 GN BH1724.  
 OS *Bacillus halodurans*.  
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;  
 OC *Bacillus*/Staphylococcus group; *Bacillus*.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RA Takami H., Nakasone K., Takaki Y.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP001512; BAB05443.1;  
 SQ SEQUENCE 16 AA; 1836 MW; 4A121EFC8BC9DBA CRC64;

Query Match 25.6%; Score 22; DB 2; Length 16;  
 Best Local Similarity 80.0%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 QGAEP 9  
 ||||  
 Db 6 QGKEP 10

## RESULT 12

ID Q9NNZ2 PRELIMINARY; PRT; 16 AA.  
 AC Q9NNZ2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE INTEGRIN ALPHA-2 SUBUNIT (FRAGMENT).  
 GN ITGA2.  
 OS *Homo sapiens* (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98421383; PubMed=9746778;  
 RA Kritzik M., Savage B., Nugent D.J., Santos S., Ruggeri Z.M.,  
 RA Kunicki T.J.;  
 RT "Nucleotide polymorphisms in the alpha2 gene define multiple alleles  
 RT that are associated with differences in platelet alpha2 beta1  
 RT density.";  
 RL Blood 92:2382-2388(1998).  
 DR EMBL; AF062039; AAF77577.1;  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1635 MW; F6A5FC64F83370A9 CRC64;

Query Match 25.6%; Score 22; DB 4; Length 16;  
 Best Local Similarity 57.1%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QROGAEP 9  
 :|||  
 Db 4 ERTGAAP 10

## RESULT 13

ID Q62207 PRELIMINARY; PRT; 11 AA.  
 AC Q62207;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE SCHWANNOMIN (FRAGMENT).  
 GN NF2.  
 OS *Mus musculus* (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; *Mus*.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=95072570; PubMed=7981675;  
 RA Huynh D.P., Nechiporuk T., Pulst S.M.;  
 RT "Alternative transcripts in the mouse neurofibromatosis type 2 (NF2)  
 RT gene are conserved and code for schwannomins with distinct C-terminal  
 RT domains.";  
 RL Hum. Mol. Genet. 3:1075-1079(1994).  
 DR EMBL; L28838; AAA57151.1;  
 DR MGD; MGI:97307; NF2.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 11 AA; 1238 MW; C51FA05774140866 CRC64;

Query Match 24.4%; Score 21; DB 11; Length 11;  
 Best Local Similarity 57.1%; Pred. No. 2.6e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QROGAEP 9  
 :|||  
 Db 2 QAQGRPP 8

## RESULT 14

ID Q9ICE5 PRELIMINARY; PRT; 12 AA.  
 AC Q9ICE5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=546CP-H7, 546CP-E4, AND 546CP-F5; TISSUE=CHOROID PLEXUS;  
RA Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;  
RT "Analysis of human immunodeficiency virus type 1 gp160 sequences from  
RT a patient with HIV dementia: evidence for monocyte trafficking into  
RT brain.";  
RL J. Neurovirol. 0:0-0(2000).  
DR EMBL; AF217155; AAF75497.1; -  
DR EMBL; AF217153; AAF75495.1; -  
DR EMBL; AF217154; AAF75496.1; -  
SQ SEQUENCE 12 AA; 1636 MW; 7ED6A2917A24005B CRC64;

Query Match 24.4%; Score 21; DB 12; Length 12;  
Best Local Similarity 66.7%; Pred. No. 2.9e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 EPRKNF 13  
| | | |  
Db 5 EIRKNY 10

RESULT 15  
Q9QV76  
ID Q9QV76 PRELIMINARY; PRT; 15 AA.  
AC Q9QV76;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE PREGNANCY ASSOCIATED MOUSE PROTEIN 1 (FRAGMENT).  
OS Mus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10095;  
RN [1]  
RP SEQUENCE.  
RA van Leuven F., Torrekens S., Van den Berghe H.;  
RL FEBS Lett. 322:219-222(1993).  
SQ SEQUENCE 15 AA; 1600 MW; 70E3C05B5D004017 CRC64;

Query Match 24.4%; Score 21; DB 11; Length 15;  
Best Local Similarity 60.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 GAEPR 10  
| | | |  
Db 6 GSEPK 10

Search completed: January 17, 2001, 13:39:13  
Job time: 178 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:37:03 ; Search time 18.64 Seconds  
(without alignments)  
15.414 Million cell updates/sec

Title: US-08-765-837-5

Perfect score: 86

Sequence: 1 ENORGAEPKRNFKVP 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 81463

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA: \*  
1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pep: \*  
2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pep: \*  
3: /cgn2.6/ptodata/2/1aa/6\_COMB.pep: \*  
4: /cgn2.6/ptodata/2/1aa/PCrUS\_COMB.pep: \*  
5: /cgn2.6/ptodata/2/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	30	34.9	8	2	US-09-016-366A-41
2	30	34.9	13	3	US-09-016-366A-41
3	30	34.9	14	1	US-08-463-620-6
4	30	34.9	14	1	US-08-463-620-7
5	30	34.9	14	2	US-08-224-917-6
6	30	34.9	14	2	US-08-224-917-7
7	30	34.9	14	2	US-08-914-853-6
8	30	34.9	14	2	US-08-914-853-7
9	30	34.9	14	4	PCT-US95-03934A-6
10	30	34.9	14	4	PCT-US95-03934A-7
11	28	32.6	15	1	US-08-477-877B-31
12	28	32.6	15	2	US-08-472-281A-31
13	28	32.6	15	2	US-08-477-989B-31
14	27	31.4	14	2	US-08-750-359-3
15	27	31.4	15	1	US-08-028-463-11
16	27	31.4	15	1	US-08-461-836-11
17	27	31.4	15	1	US-08-469-526A-195
18	27	31.4	15	2	US-08-734-591A-185
19	27	31.4	15	3	US-08-470-335-185
20	26	30.2	9	3	US-08-891-271-19
21	26	30.2	11	2	US-08-633-148-15
22	26	30.2	13	3	US-09-052-830B-5
23	26	30.2	14	2	US-08-637-759B-110
24	26	30.2	14	3	US-08-871-355A-110
25	26	30.2	15	2	US-08-685-357B-2
26	26	30.2	15	2	US-08-392-338A-5
27	26	30.2	15	3	US-09-166-750-5
28	26	30.2	15	3	US-09-166-093-5

29 26 30.2 15 3 US-09-172-019-5  
30 26 30.2 15 3 US-09-166-094-5  
31 26 30.2 16 1 US-07-995-657-3  
32 26 30.2 16 1 US-08-474-587-3  
33 26 30.2 16 2 US-08-493-594-7  
34 25 29.1 13 1 US-08-149-809-7  
35 25 29.1 13 1 US-08-520-599-4  
36 25 29.1 13 2 US-08-572-951-19  
37 25 29.1 13 2 US-08-572-951-20  
38 25 29.1 13 2 US-08-572-951-40  
39 25 29.1 13 2 US-08-878-844-1  
40 25 29.1 13 3 US-08-409-901-1  
41 25 29.1 14 1 US-08-357-264-7  
42 25 29.1 14 1 US-08-357-264-8  
43 25 29.1 14 1 US-08-672-514-7  
44 25 29.1 14 1 US-08-672-514-8  
45 25 29.1 14 2 US-08-878-844-3

#### ALIGNMENTS

RESULT 1  
US-09-016-366A-41  
; Sequence 41, Application US/09016366A  
; Patent No. 5955431  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; APPLICANT: Huang, Chifu  
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,366A  
; FILING DATE: January 30, 1998  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/037,090  
; FILING DATE: 05-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7093  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 41:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-016-366A-41

Query Match 34.9%; Score 30; DB 2; Length 8;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gap 0;

Qy 3 ORQGAEPK 10  
I : I I I I  
Db 1 QKRGREPR 8

## RESULT 2

US-09-273-565-68  
; Sequence 68, Application US/09273565A  
; Patent No. 6166190  
; GENERAL INFORMATION:  
; APPLICANT: FUJIWARA, TSUTOMU  
; APPLICANT: WATANABE, TAKESHI  
; APPLICANT: HORIE, MASATO  
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN  
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME  
; FILE REFERENCE: Q-53599  
; CURRENT APPLICATION NUMBER: US/09/273,565A  
; CURRENT FILING DATE: 1999-03-22  
; EARLIER APPLICATION NUMBER: 09/055,699  
; EARLIER FILING DATE: 1998-04-07  
; EARLIER APPLICATION NUMBER: 08/820,170  
; EARLIER FILING DATE: 1997-03-19  
; EARLIER APPLICATION NUMBER: JP 63410/1996  
; EARLIER FILING DATE: 1996-03-19  
; EARLIER APPLICATION NUMBER: JP 69163/1997  
; EARLIER FILING DATE: 1997-03-05  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 68  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Bovine sp.  
US-09-273-565-68

Query Match 34.9%; Score 30; DB 3; Length 13;  
Best Local Similarity 46.2%; Pred. No. 31;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 ORQGAEPKRNFKV 15  
I : I I I : I I  
Db 1 OSQGLSPAQAFK 13

## RESULT 3

US-08-463-620-6  
; Sequence 6, Application US/08463620  
; Patent No. 5789216  
; GENERAL INFORMATION:  
; APPLICANT: Lou, Lillian Lien-Li  
; APPLICANT: Barnett, Jimmy Wayne  
; TITLE OF INVENTION: Cloning and Expression of Human GMP  
; TITLE OF INVENTION: Cloning and Expression of Inhibitors of  
; TITLE OF INVENTION: Synthetase, its use in Screening for Inhibitors of  
; TITLE OF INVENTION: Human  
; TITLE OF INVENTION: Human  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Syntex (USA) Inc.  
; STREET: 3401 Hillview Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,620  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/224,917  
; FILING DATE: 08-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Peries, Rohan  
; REGISTRATION NUMBER: 35,752  
; REFERENCE/DOCKET NUMBER: 28060  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)-852-1698  
; TELEFAX: (415)-496-3529  
; INFORMATION FOR SEQ ID NO: 1  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

; APPLICATION NUMBER: US 08/224,917  
; FILING DATE: 08-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Peries, Rohan  
; REGISTRATION NUMBER: 35,752  
; REFERENCE/DOCKET NUMBER: 28060  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)-852-1698  
; TELEFAX: (415)-496-3529  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: internal  
; US-08-463-620-6

Query Match 34.9%; Score 30; DB 1; Length 14;  
Best Local Similarity 62.5%; Pred. No. 34;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 8 EPRKNFKV 15  
I : I I I : I I  
Db 3 EPLKDFIK 10

## RESULT 4

US-08-463-620-7  
; Sequence 7, Application US/08463620  
; Patent No. 5789216  
; GENERAL INFORMATION:  
; APPLICANT: Lou, Lillian Lien-Li  
; APPLICANT: Barnett, Jimmy Wayne  
; TITLE OF INVENTION: Cloning and Expression of Human GMP  
; TITLE OF INVENTION: Cloning and Expression of Inhibitors of  
; TITLE OF INVENTION: Synthetase, its use in Screening for Inhibitors of  
; TITLE OF INVENTION: Human  
; TITLE OF INVENTION: Human  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Syntex (USA) Inc.  
; STREET: 3401 Hillview Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,620  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/224,917  
; FILING DATE: 08-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Peries, Rohan  
; REGISTRATION NUMBER: 35,752  
; REFERENCE/DOCKET NUMBER: 28060  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)-852-1698  
; TELEFAX: (415)-496-3529  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
US-08-463-620-7

Query Match 34.9%; Score 30; DB 1; Length 14;  
Best Local Similarity 62.5%; Pred. No. 34;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 EPRKNFVK 15  
|||:|:|  
Db 3 EPLKDFIK 10

## RESULT 5

US-08-224-917-6  
Sequence 6, Application US/08224917  
Patent No. 5965350

## GENERAL INFORMATION:

APPLICANT: Lou, Lillian Lien-Li  
TITLE OF INVENTION: Cloning and Expression of Human GMP  
TITLE OF INVENTION: Synthetase, its use in Screening for Inhibitors of Human  
TITLE OF INVENTION: GMP Synthetase and Inhibitors of Human GMP Synthetase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Syntex (USA) Inc.  
STREET: 3401 Hillview Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/224,917  
FILING DATE: 08-APR-1994  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Perles, Rohan  
REGISTRATION NUMBER: 35,752  
REFERENCE/DOCKET NUMBER: 28060  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)-852-1698

TELEFAX: (415)-496-3529  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
US-08-224-917-6

Query Match 34.9%; Score 30; DB 2; Length 14;  
Best Local Similarity 62.5%; Pred. No. 34;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 EPRKNFVK 15  
|||:|:|  
Db 3 EPLKDFIK 10

## RESULT 6

US-08-224-917-7

Sequence 7, Application US/08224917  
Patent No. 5965350  
GENERAL INFORMATION:

APPLICANT: Lou, Lillian Lien-Li  
TITLE OF INVENTION: Cloning and Expression of Human GMP  
TITLE OF INVENTION: Synthetase, its use in Screening for Inhibitors of Human  
TITLE OF INVENTION: GMP Synthetase and Inhibitors of Human GMP Synthetase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Syntex (USA) Inc.  
STREET: 3401 Hillview Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/224,917  
FILING DATE: 08-APR-1994  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Perles, Rohan  
REGISTRATION NUMBER: 35,752  
REFERENCE/DOCKET NUMBER: 28060  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)-852-1698

TELEFAX: (415)-496-3529  
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO

FRAGMENT TYPE: internal  
US-08-224-917-7

Query Match 34.9%; Score 30; DB 2; Length 14;  
Best Local Similarity 62.5%; Pred. No. 34;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 EPRKNFVK 15  
|||:|:|  
Db 3 EPLKDFIK 10

## RESULT 7

US-08-914-853-6

Sequence 6, Application US/08914853  
Patent No. 5998186  
GENERAL INFORMATION:

APPLICANT: Lou, Lillian Lien-Li  
TITLE OF INVENTION: Cloning and Expression of Human GMP  
TITLE OF INVENTION: Synthetase, its use in Screening for Inhibitors of  
TITLE OF INVENTION: Human  
TITLE OF INVENTION: GMP Synthetase and Inhibitors of Human GMP Synthetase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Syntex (USA) Inc.  
STREET: 3401 Hillview Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:





COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03934A  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: Internal  
PCT-US95-03934A-7

Query Match 34.9%; Score 30; DB 4; Length 14;  
Best Local Similarity 62.5%; Pred. No. 34;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 EPRKNFVK 15  
|||1:1:1  
DB 3 EPLKDFIK 10

RESULT 11  
US-08-477-877B-31  
; Sequence 31, Application US/0847787B  
; Patent No. 5730979  
; GENERAL INFORMATION:  
; APPLICANT: Bazin, Herv  
; APPLICANT: Latinne, Dominique  
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation  
; NUMBER OF SEQUENCES: 96  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
; ADDRESSEE: Cecchi, Stewart & Olstein  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,877B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/407,009  
; FILING DATE: 29-MAR-1995  
; APPLICATION NUMBER: 08/119,032  
; FILING DATE: 09-SEP-1993  
; APPLICATION NUMBER: 08/027,008  
; FILING DATE: 05-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Olstein, Elliot M.  
; REGISTRATION NUMBER: 24,025  
; REFERENCE/DOCKET NUMBER: 61750-146  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide

US-08-477-877B-31

Query Match 32.6%; Score 28; DB 1; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 QROGAEPK 11  
|||1:1:1  
DB 4 QRPQSPRR 12

RESULT 12

US-08-472-281A-31  
; Sequence 31, Application US/08472281A  
; Patent No. 5817311  
; GENERAL INFORMATION:  
; APPLICANT: Bazin, Herv  
; APPLICANT: Latinne, Dominique  
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation  
; NUMBER OF SEQUENCES: 96  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
; ADDRESSEE: Cecchi, Stewart & Olstein  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,281A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/407,009  
; FILING DATE: 29-MAR-1995  
; APPLICATION NUMBER: 08/119,032  
; FILING DATE: 09-SEP-1993  
; APPLICATION NUMBER: 08/027,008  
; FILING DATE: 05-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Olstein, Elliot M.  
; REGISTRATION NUMBER: 24,025  
; REFERENCE/DOCKET NUMBER: 61750-142  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
US-08-472-281A-31

Query Match 32.6%; Score 28; DB 2; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 QROGAEPK 11  
|||1:1:1  
DB 4 QRPQSPRR 12

RESULT 13

US-08-477-989B-31

; Sequence 31, Application US/08477989B  
; Patent No. 5951983  
; GENERAL INFORMATION:  
; APPLICANT: Bazin, Herv  
; APPLICANT: Latanne, Dominique  
; APPLICANT: Kaplan, Ruth  
; APPLICANT: Kieber-Emmons, Thomas  
; APPLICANT: Postema, Christina E.  
; APPLICANT: White-Scharf, Mary  
; TITLE OF INVENTION: LO-CD2a Antibody and Uses  
; TITLE OF INVENTION: Thereof for Inhibiting  
; TITLE OF INVENTION: T-Cell Activation and  
; TITLE OF INVENTION: Proliferation  
; NUMBER OF SEQUENCES: 96  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
; ADDRESSEE: Cecchi, Stewart & Olstein  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07068

## COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,989B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:

## APPLICATION INFORMATION:

; APPLICATION NUMBER: 08/407,009  
; FILING DATE: 29-MAR-1995  
; APPLICATION NUMBER: 08/119,032  
; FILING DATE: 09-SEP-1993  
; APPLICATION NUMBER: 08/027,008  
; FILING DATE: 05-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Olstein, Elliot M.  
; REGISTRATION NUMBER: 24,025  
; REFERENCE/DOCKET NUMBER: 61750-147  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744

## INFORMATION FOR SEQ ID NO: 31:

; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear

## MOLECULE TYPE: polypeptide

; US-08-477-989B-31

Query Match 32.6%; Score 28; DB 2; Length 15;  
Best Local Similarity 55.6%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 QROGAEPK 11

Db 4 QROGQSPRR 12

## RESULT 14

US-08-750-359-3  
; Sequence 3, Application US/08750359  
; Patent No. 5916561  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Monoclonal antibody against CD44v6  
; NUMBER OF SEQUENCES: 6  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/750,359  
; FILING DATE:  
; CLASSIFICATION: 424

## INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single stranded  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
; US-08-750-359-3

Query Match 31.4%; Score 27; DB 2; Length 14;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENORQGAEP 9

Db 4 ENEWQGNP 12

## RESULT 15

US-08-028-463-11  
; Sequence 11, Application US/08028463  
; Patent No. 5731176  
; GENERAL INFORMATION:

## APPLICATION INFORMATION:

; APPLICANT: BEPPU, TERUHIKO  
; APPLICANT: YAMADA, HIDEAKI  
; APPLICANT: NAGASAWA, TORU  
; APPLICANT: HORINOCHI, SUCHAYU  
; APPLICANT: NISHIVAMA, MAKATO

## TITLE OF INVENTION: DNA FRAGMENT ENCODING A POLYPEPTIDE

## TITLE OF INVENTION: HAVING NITRILE HYDRATASE ACTIVITY, A TRANSFORMANT CONTAININ

## TITLE OF INVENTION: GENE AND A PROCESS FOR THE PRODUCTION OF AMIDES USING THE

## TITLE OF INVENTION: TRANSFORMANT

## NUMBER OF SEQUENCES: 15

## CORRESPONDENCE ADDRESS:

; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 AVENUE OF THE AMERICAS

; CITY: NEW YORK

; STATE: N.Y.

; COUNTRY: U.S.A.

; ZIP: 10036

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/028,463

; FILING DATE: 09-MAR-1993

## CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/694,747

; FILING DATE: 02-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: MISROCK, S. LESLIE

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 7005-023-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 790-9090

; TELEFAX: 212 869-9741

## INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 15 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

;  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Rhodococcus rhodochrous  
; STRAIN: J-1 (FERM BP-1478)  
US-08-028-463-11

Query Match 31.4%; Score 27; DB 1; Length 15;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 NORQGAEPRKN 12  
| | | | |  
Db 4 NPVQGTLPXN 14

Search completed: January 17, 2001, 13:37:04  
Job time: 124 sec



GenCore version 4.5  
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 17, 2001, 13:25:20 ; Search time 15.26 Seconds  
(without alignments)  
29.130 Million cell updates/sec

Title: US-08-765-837-1

Perfect score: 66

Sequence: 1 RTTQSDQEIDY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 92618

Minimum DB seq length: 0

Maximum DB seq length: 13

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_36:\*
- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:\*
  - 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:\*
  - 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:\*
  - 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:\*
  - 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:\*
  - 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:\*
  - 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:\*
  - 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:\*
  - 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:\*
  - 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:\*
  - 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:\*
  - 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:\*
  - 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:\*
  - 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:\*
  - 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:\*
  - 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:\*
  - 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:\*
  - 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:\*
  - 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:\*
  - 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:\*
  - 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	13	17	R90945
2	29	43.9	13	17	R28967
3	26	39.4	13	17	W06271
4	25	37.9	12	7	P05552
5	24	36.4	10	20	Y07173
6	24	36.4	10	20	Y07131
7	24	36.4	11	16	R82679
8	24	36.4	13	20	Y27136
9	23	34.8	7	15	R46053
10	23	34.8	7	15	R46057
11	23	34.8	8	14	R43245
12	23	34.8	8	14	R43246

13	23	34.8	8	16	R78840	P. falciparum circ
14	23	34.8	8	19	W54663	Peptide from P.fal
15	23	34.8	8	19	W54664	Peptide from P.fal
16	23	34.8	8	20	Y03662	Amino acid sequenc
17	23	34.8	10	19	W79734	Peptide p5-. Synt
18	23	34.8	12	13	R27122	Hirudin analogue of
19	23	34.8	13	9	P82670	Peptide capable of
20	23	34.8	13	17	W06267	ITIM motif used to
21	23	34.8	13	17	W06269	ITIM motif used to
22	23	34.8	13	17	R92853	Cell adhesion modifi
23	23	34.8	13	19	W54540	Synthetic peptide
24	23	34.8	13	19	W54541	Synthetic peptide
25	23	34.8	13	20	W82749	Bovine aprotinin w
26	23	34.8	13	20	W82750	Bovine aprotinin w
27	23	34.8	13	20	W82751	Bovine aprotinin w
28	22	33.3	6	15	R62208	CENP-B protein and
29	22	33.3	7	15	R50265	Peptide correspond
30	22	33.3	8	20	Y48741	Membrane dipeptida
31	22	33.3	9	16	R87291	Plasmodium falcipa
32	22	33.3	9	20	Y25216	HIV gag protein de
33	22	33.3	9	20	W84220	Antigenic peptide
34	22	33.3	9	21	Y90473	Yeast Smi C-termi
35	22	33.3	10	9	P82330	E.coli adventitia
36	22	33.3	10	16	R74936	H-CDR-3 of anti-id
37	22	33.3	10	21	Y77821	Human BRCA1 protei
38	22	33.3	10	21	Y77822	Rhesus BRCA1 protei
39	22	33.3	12	14	R44862	Inhibitor of TK of
40	22	33.3	12	16	R62589	Human hepatitis C
41	22	33.3	12	20	Y13405	HSV specific anti
42	22	33.3	12	20	Y10214	T cell epitope/MHC
43	22	33.3	13	18	W24966	C3/C4/C5-derived c
44	21	31.8	7	15	R62118	Hydrophilic, acidi
45	21	31.8	7	15	W21259	Hydroxymethylalula

ALIGNMENTS

RESULT 1

- R90945 ID R90945 standard; peptide; 13 AA.  
AC R90945;  
DT 09-SEP-1996 (first entry)  
DE Factor VIII antigenic peptide corresp. to residues Arg1652-Tyr1664.  
KW Factor VIII; modification; inhibitor activity; binding; antibody;  
KW von Willebrand factor; immune disorder.  
OS Synthetic.  
XX WO9602572-A2.  
XX 01-FEB-1996.  
XX 14-JUL-1995; 95WO-BE00068.  
XX 14-JUL-1994; 94BE-0000666.  
XX (CROI-) CROIX ROUGE BELGQUE.  
XX Di Giambattista M, Laub R;  
XX WPI; 1996-105861/11.  
XX Factor VIII antigenic polypeptide fragments and epitope(s) - also  
XX inhibitors of factor VIII and anti-inhibitors, useful for e.g.  
XX preventing and treating immune disorders involving inhibition of  
XX factor VIII binding  
XX Claim 8; Page 30; 45pp; French.

I closed  
only hit (1/1 hits)

XX Peptides R90945-64 are derived from the factor VIII protein, esp. from a  
 CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,  
 CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII  
 CC and derived peptides can be used to prevent the activity of inhibitors of  
 CC factor VIII binding to von Willebrand factor, esp. antibodies, thus  
 CC preventing or treating immune disorders.  
 XX  
 SQ Sequence 13 AA;

Query Match 100.0%; Score 66; DB 17; Length 13;  
 Best Local Similarity 100.0%; Pred. NO. 6.4e-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RTTLQSDQEEIDY 13  
 Db 1 rttlqsdqeidy 13

RESULT 2  
 R28967  
 ID R28967 standard; peptide; 13 AA.  
 XX  
 AC R28967;  
 XX  
 XX 23-MAR-1993 (first entry)  
 DT  
 XX  
 DE Selectin based anti-adhesion peptide.

XX ELAM; LHR; pertussis toxin; PT; GMPI40; metastasis; coagulation;  
 KW inflammation; vaccine.  
 XX  
 XX Synthetic.  
 OS  
 XX WO9219646-A.  
 PN  
 XX 12-NOV-1992.  
 PD  
 XX 04-MAY-1992; 92WO-US03701.  
 PF  
 XX 03-MAY-1991; 91US-0695532.  
 PR  
 XX (UVRQ ) UNIV ROCKEFELLER.  
 PA  
 XX Masure HR, Tuomanen E;  
 PI  
 XX WPI; 1992-398798/48.  
 DR  
 XX

Peptide(s) inhibiting adhesion between circulating cells and  
 endothelia - contain sequences homologous to Bordetella pertussis  
 PT and selectins 1, used for treating malignancy or inflammation  
 PT caused by infection or auto-immune disease  
 XX  
 XX Claim 4; Page 52; 74pp; English.

XX The synthetic peptide has a sequence based on regions of homology  
 CC between Bordetella pertussis toxin S-subunits. The peptide acts  
 CC as an analogue of the selectins ELAM, GMP140 and hHR. Antibodies  
 CC to this peptide can block the function of selectins and prevent the  
 CC adherence of circulating cells to endothelial cells and inhibit  
 CC inflammatory, metastatic or coagulation processes. The peptide also  
 CC functions as a carrier to direct selected therapeutic agents to  
 CC cells bearing selectin-specific receptors such as leukocytes,  
 CC platelets or high endothelial venules. The antibodies will further  
 CC prevent binding of B. pertussis to eukaryotic cells such as ciliated  
 CC cells and macrophages to interfere with the persistence of B.  
 CC pertussis infection of a host. In addition, the peptide can function  
 CC as a vaccine to protect against B. pertussis infection.  
 CC See also R28965-78.

XX Sequence 13 AA;  
 SQ

Query Match 43.9%; Score 29; DB 13; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 TTLQSDQEEIDY 13  
 Db 1 tvlaiqneeley 12

RESULT 3  
 W06271  
 ID W06271 standard; peptide; 13 AA.  
 XX  
 AC W06271;  
 XX  
 XX 15-MAY-1997 (first entry)  
 DT  
 XX  
 DE ITIM motif used to regulate haematopoietic cell function.

XX ITIM; immunoreceptor tyrosine inhibitory motif; haematopoiesis;  
 KW blood; clotting; antibody response; inflammation; cytokine release;  
 KW immunoproliferative; immunodeficiency; autoimmune; infectious;  
 KW disease; cancer; allergy; drug screening.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 9 /note= "preferably non-hydrolysable  
 FT phosphotyrosine analogue"  
 FT  
 XX W09626961-A2.  
 PN  
 XX 06-SEP-1996.  
 PD  
 XX 28-FEB-1996; 96WO-US02289.  
 PF  
 XX 28-FEB-1995; 95US-0397628.  
 PR  
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.  
 PA  
 XX Cambier JC;  
 PI  
 XX WPI; 1996-412740/41.  
 DR  
 XX  
 XX New immuno-receptor tyrosine-inhibitory motif peptide(s) - used to  
 PT develop methods for regulating inflammatory and immune responses and  
 PT for drug screening  
 PT  
 XX Claim 5; Page 88; 99pp; English.

XX W06255-W06276 are immunoreceptor tyrosine-inhibitory motifs (ITIMs) that  
 CC bind to protein tyrosine phosphatase 1C (PTPIC), PTPID, p160 or p70 and  
 CC in doing so, inhibit signal transduction pathways involving these  
 CC molecules. ITIM motif containing peptides can be used for regulating  
 CC inflammatory and immune responses and for drug screening. The peptides  
 CC can also be used to regulate haematopoietic cell function, especially  
 CC antibody responses, blood clotting initiation, inflammatory responses,  
 CC release of cytokines or mediators, or cytotoxicity and are useful for  
 CC treating an immunoproliferative, immunodeficiency, autoimmune or  
 CC infectious disease, cancer or allergic response. The peptides can also  
 CC be used in the production of antibodies and in assays and diagnostics.  
 XX  
 SQ Sequence 13 AA;

Query Match 39.4%; Score 26; DB 17; Length 13;  
 Best Local Similarity 44.4%; Pred. NO. 1e+02;  
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 QSDQEEIDY 13  
 : | | | |

```

Db      1 eqdpeevty 9

RESULT 4
P60552
ID      P60552 standard; Protein; 12 AA.
XX
AC      P60552;
XX
DT      24-JUN-1991 (first entry)
XX
DE      Fragment of Sall rotavirus major outer shell glycoprotein.
XX
KW      Hu/Australia/5/77 rotavirus; vaccine.
XX
OS      Sall simian rotavirus.
XX
PN      W08505122-A.
XX
PD      21-NOV-1985.
XX
PF      29-APR-1985; 85WO-AU00096.
XX
PR      27-APR-1984; 84AU-0004733.
XX
PR      29-APR-1985; 85WO-AU00096.
XX
PR      01-JAN-1985; 85AU-0042970.
XX
PA      (UYME-) UNIV OF MELBOURNE.
XX
PI      Holmes I, Dyall-Smith ML;
XX
DR      WPI; 1986-028178/04.
XX
PT      RNA gene segment coding for outer capsid glyco:protein of
PT      rotavirus - useful in expression of antigenic viral proteins by
PT      bacteria for use in vaccines and diagnostic prods.
XX
PS      Disclosure; Fig 5; 24pp; English.
XX
CC      The protein fragment was isolated for comparison with the rotavirus
CC      segment 8 protein fragment isolated from the Hu/Australia/5/11.
CC      The segment 8 product and fragments comprise at least part of
CC      the major outer capsid glycoprotein, they may be expressed from a
CC      transformed host and are useful as antigens for vaccination and
CC      diagnosis of the rotavirus.
XX
SQ      Sequence 12 AA;

Query Match      37.9%; Score 25; DB 7; Length 12;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      3 TLQSDQEEI 11
Db      4 tlqldmse 12

RESULT 5
Y07173
ID      Y07173 standard; Peptide; 10 AA.
XX
AC      Y07173;
XX
DT      02-JUL-1999 (first entry)
XX
DE      HLA binding peptide.
XX
KW      Cancer associated antigen; diagnosis; research; treatment; human;
KW      breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW      prostate cancer.
XX
OS      Homo sapiens.
XX

Db      1 eqdpeevty 9

RESULT 4
P60552
ID      P60552 standard; Protein; 12 AA.
XX
AC      P60552;
XX
DT      24-JUN-1991 (first entry)
XX
DE      Fragment of Sall rotavirus major outer shell glycoprotein.
XX
KW      Hu/Australia/5/77 rotavirus; vaccine.
XX
OS      Sall simian rotavirus.
XX
PN      W08505122-A.
XX
PD      21-NOV-1985.
XX
PF      29-APR-1985; 85WO-AU00096.
XX
PR      27-APR-1984; 84AU-0004733.
XX
PR      29-APR-1985; 85WO-AU00096.
XX
PR      01-JAN-1985; 85AU-0042970.
XX
PA      (UYME-) UNIV OF MELBOURNE.
XX
PI      Holmes I, Dyall-Smith ML;
XX
DR      WPI; 1986-028178/04.
XX
PT      RNA gene segment coding for outer capsid glyco:protein of
PT      rotavirus - useful in expression of antigenic viral proteins by
PT      bacteria for use in vaccines and diagnostic prods.
XX
PS      Disclosure; Fig 5; 24pp; English.
XX
CC      The protein fragment was isolated for comparison with the rotavirus
CC      segment 8 protein fragment isolated from the Hu/Australia/5/11.
CC      The segment 8 product and fragments comprise at least part of
CC      the major outer capsid glycoprotein, they may be expressed from a
CC      transformed host and are useful as antigens for vaccination and
CC      diagnosis of the rotavirus.
XX
SQ      Sequence 12 AA;

Query Match      37.9%; Score 25; DB 7; Length 12;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      3 TLQSDQEEI 11
Db      4 tlqldmse 12

RESULT 5
Y07173
ID      Y07173 standard; Peptide; 10 AA.
XX
AC      Y07173;
XX
DT      02-JUL-1999 (first entry)
XX
DE      HLA binding peptide.
XX
KW      Cancer associated antigen; diagnosis; research; treatment; human;
KW      breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW      prostate cancer.
XX
OS      Homo sapiens.
XX

XX
PN      W09904265-A2.
XX
PD      28-JAN-1999.
XX
PF      15-JUL-1998; 98WO-US14679.
XX
PR      22-JUN-1998; 98US-0102322.
PR      17-JUL-1997; 97US-0896164.
PR      10-OCT-1997; 97US-0061599.
PR      10-OCT-1997; 97US-0061765.
PR      10-OCT-1997; 97US-0948705.
PR      11-OCT-1997; 97GB-0021697.
XX
PA      (LUDW-) LUDWIG INST CANCER RES.
XX
PI      Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI      Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
PI      Tureci O;
XX
DR      WPI; 1999-132448/11.
XX
PT      New isolated cancer associated nucleic acids and polypeptides -
PT      isolated using sera from cancer patients, used to develop products
PT      for the diagnosis, monitoring or treatment of cancers
XX
PS      Example 16; Page 763; 787pp; English.
XX
CC      The invention relates to a method for diagnosing a disorder characterised
CC      by expression of a human cancer associated antigen precursor coded for by
CC      a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC      biological sample isolated from a subject with an agent that specifically
CC      binds to the NAM, an expression product or a fragment of an expression
CC      product complexed with an HLA molecule; and (b) determining the
CC      interaction between the agent and the NAM or the expression product as a
CC      determination of the disorder. The products and methods can be used in
CC      the diagnosis, monitoring, research, or treatment of conditions
CC      characterised by the expression of various cancer associated antigens.
CC      The invention provides nucleic acid sequences and encoded polypeptides
CC      which are cancer associated antigen precursors expressed in human breast
CC      cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC      lung cancer.
XX
SQ      Sequence 10 AA;

Query Match      36.4%; Score 24; DB 20; Length 10;
Best Local Similarity 56.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      7 DQEEID 12
Db      4 dkedid 9

RESULT 6
Y07131
ID      Y07131 standard; Peptide; 10 AA.
XX
AC      Y07131;
XX
DT      02-JUL-1999 (first entry)
XX
DE      HLA binding peptide.
XX
KW      Cancer associated antigen; diagnosis; research; treatment; human;
KW      breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW      prostate cancer.
XX
OS      Homo sapiens.
XX
PN      W09904265-A2.
XX

```

```

PD XX 28-JAN-1999.
PF XX
XX XX 15-JUL-1998; 98WO-US14679.
XX XX
XX XX 22-JUN-1998; 98US-0102322.
XX XX 17-JUL-1997; 97US-0896164.
XX XX 10-OCT-1997; 97US-0061599.
XX XX 10-OCT-1997; 97US-0061765.
XX XX 10-OCT-1997; 97US-0948705.
XX XX 11-OCT-1997; 97GB-0021697.
XX XX
XX XX (LUDW-) LUDWIG INST CANCER RES.
XX XX
XX XX Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
XX XX Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
XX XX Tureci O;
XX XX
XX XX WPI; 1999-132448/11.
XX XX
XX XX New isolated cancer associated nucleic acids and polypeptides -
XX XX isolated using sera from cancer patients, used to develop products
XX XX for the diagnosis, monitoring or treatment of cancers
XX XX
XX XX Example 16; Page 756; 787pp; English.
XX XX
XX XX The invention relates to a method for diagnosing a disorder characterised
XX XX by expression of a human cancer associated antigen precursor coded for by
XX XX a nucleic acid molecule (NAM). The method comprises: (a) contacting a
XX XX biological sample isolated from a subject with an agent that specifically
XX XX binds to the NAM, an expression product or a fragment of an expression
XX XX product complexed with an HLA molecule; and (b) determining the
XX XX interaction between the agent and the NAM or the expression product as a
XX XX determination of the disorder. The products and methods can be used in
XX XX the diagnosis, monitoring, research, or treatment of conditions
XX XX characterised by the expression of various cancer associated antigens.
XX XX The invention provides nucleic acid sequences and encoded polypeptides
XX XX which are cancer associated antigen precursors expressed in human breast
XX XX cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
XX XX lung cancer.
XX XX
XX XX Sequence 10 AA;
SQ

```

```

Query Match 36.4%; Score 24; DB 20; Length 10;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 DQEEID 12
   |::|
Db 4 dkdid 9

```

```

RESULT 7
R82679
ID R82679 standard; Protein; 11 AA.
XX
XX R82679;
XX
XX 23-JAN-1996 (first entry)
XX
XX V8 fragment of CD45 associating protein pp32.
XX
XX CD45; immune system; disease; pp32; T-lymphocyte; T-cell;
XX activation; isoform; CD2; phorbol ester; identification; screening;
XX monoclonal antibody.
XX
XX Homo sapiens.
XX
XX WO9521916-A2.
XX
XX 17-AUG-1995.
XX
XX 09-FEB-1995; 95WO-US01618.

```

```

XX 14-FEB-1994; 94US-0197793.
XX
XX (BADI ) BASF AG.
XX
XX Meuer S, Ratnofsky S, Schoenhaut D, Schraven B;
XX
XX WPI; 1995-293118/38.
XX N-PSDB; T03910.
XX
XX Nucleic acid encoding a human pp32 protein which associates with
XX CD45 - useful for treatment of immune system diseases and in the
XX identification of T cell activation inhibitors
XX
XX Example 9; Page 25; 49pp; English.
XX
XX pp32 is specifically associated with, and is a potential substrate of
XX CD45. pp32 is constitutively phosphorylated on serine in resting T-
XX cells. In such cells it exists in two isoforms, pp32 low and pp32
XX high. Both isoforms show rapid changes during T-cell activation.
XX These changes take place within five minutes after stimulation of T-
XX lymphocytes with monoclonal antibodies specific for CD2 or with
XX phorbol esters. pp32 is useful for treating diseases of the immune
XX system and to identify inhibitors of T-cell activation. Peptide
XX fragments of pp32 (R82673-81) were used to produce degenerate
XX primers (T03911-19) which were then used to produce a probe to
XX isolate a full length pp32 cDNA clone.
XX
XX Sequence 11 AA;
SQ

```

```

Query Match 36.4%; Score 24; DB 16; Length 11;
Best Local Similarity 42.9%; Pred. No. 1.9e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 7 DQEEID 13
   |::|
Db 2 deqtdy 8

```

```

RESULT 8
Y27136
ID Y27136 standard; peptide; 13 AA.
XX
XX Y27136;
XX
XX 15-SEP-1999 (first entry)
XX
XX Human munc13 (Hmunc13) polypeptide N-terminal segment.
XX
XX Hmunc13; glucose-regulated; kidney; renal; microvascular; hyperglycemia;
XX diabetes mellitus; apoptosis; vesicle trafficking; glomerulopathy;
XX renal failure; human.
XX
XX Homo sapiens.
XX
XX WO9931134-A1.
XX
XX 24-JUN-1999.
XX
XX 19-NOV-1998; 98WO-CA01061.
XX
XX 12-DEC-1997; 97US-0069352.
XX
XX (SILV/) SILVERMAN M.
XX (SONG/) SONG Y.
XX
XX Silverman M, Song Y;
XX
XX WPI; 1999-418755/35.
XX
XX A new glucose-regulated human munc13 gene useful for treating
XX insulin dependent and independent diabetes mellitus
XX

```





XX 04-MAY-1994 (first entry)  
 DT Cytotoxic T lymphocyte recognition/induction peptide.  
 DE CTL; vaccine; malaria; specific antigen-derived.  
 XX Synthetic.  
 OS WO9320103-A.  
 PN 14-OCT-1993.  
 PD 05-APR-1993; 93WO-GB00711.  
 XX 03-APR-1992; 92GB-0008068.  
 PR 20-AUG-1992; 92GB-0017704.  
 XX (ISIS-) ISIS INNOVATION LTD.  
 PA Elvin J, Gotch FM, Hill AV, McMichael AJ, Whittle HC;  
 PI WPI; 1993-336833/42.  
 DR Peptide(s) recognising or inducing cytotoxic T lymphocytes -  
 XX useful in vaccines against malaria or HIV-2, derived from  
 PT specific antigen and human leukocyte antigen contg. class I  
 PT restricted epitope  
 XX Claim 6; Page 30; 35pp; English.  
 PS The sequence is that of peptide cp26 which is recognised by, or can  
 CC induce, cytotoxic T lymphocytes. It may be useful in vaccines against  
 CC malaria.  
 CC Sequence 8 AA;  
 SQ Query Match 34.8%; Score 23; DB 14; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QOEIDY 13  
 DB 3 kdely 8

RESULT 12  
 R43246  
 ID R43246 standard; peptide; 8 AA.  
 AC R43246;  
 XX 04-MAY-1994 (first entry)  
 DT Cytotoxic T lymphocyte recognition/induction peptide.  
 DE CTL; vaccine; malaria; specific antigen-derived.  
 XX Synthetic.  
 OS WO9320103-A.  
 PN 14-OCT-1993.  
 PD 05-APR-1993; 93WO-GB00711.  
 XX 03-APR-1992; 92GB-0008068.  
 PR 20-AUG-1992; 92GB-0017704.  
 XX (ISIS-) ISIS INNOVATION LTD.  
 PA Elvin J, Gotch FM, Hill AV, McMichael AJ, Whittle HC;  
 PI WPI; 1993-336833/42.  
 DR Peptide(s) recognising or inducing cytotoxic T lymphocytes -  
 XX useful in vaccines against malaria or HIV-2, derived from  
 PT specific antigen and human leukocyte antigen contg. class I  
 PT restricted epitope  
 XX Claim 6; Page 30; 35pp; English.  
 PS The sequence is that of peptide cp26 which is recognised by, or can  
 CC induce, cytotoxic T lymphocytes. It may be useful in vaccines against  
 CC malaria.  
 CC Sequence 8 AA;  
 SQ Query Match 34.8%; Score 23; DB 14; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

XX WPI; 1993-336833/42.  
 DR Peptide(s) recognising or inducing cytotoxic T lymphocytes -  
 XX useful in vaccines against malaria or HIV-2, derived from  
 PT specific antigen and human leukocyte antigen contg. class I  
 PT restricted epitope  
 XX Claim 6; Page 30; 35pp; English.  
 PS The sequence is that of peptide cp29 which is recognised by, or can  
 CC induce, cytotoxic T lymphocytes. It may be useful in vaccines against  
 CC malaria.  
 CC Sequence 8 AA;  
 SQ Query Match 34.8%; Score 23; DB 14; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QOEIDY 13  
 DB 3 kdely 8

RESULT 13  
 R78840  
 ID R78840 standard; peptide; 8 AA.  
 AC R78840;  
 XX 27-MAR-1996 (first entry)  
 DT P. falciparum circumsp. 368-375 cytotoxic T lymphocyte epitope.  
 XX Circumsp. 368-375; cytotoxic T; CTL; epitope: helper T; HTL; cell;  
 KW lymphocyte; viruses; parasites; tumours; antigens; treatment;  
 KW disease prevention.  
 XX Plasmodium falciparum.  
 OS WO9522317-A1.  
 PN 24-AUG-1995.  
 XX 16-FEB-1995; 95WO-US02121.  
 PF 16-FEB-1994; 94US-0197484.  
 PR (CYTE-) CYTEL CORP.  
 PA Ceut RW, Grey H, Sette AD, Vitiello MA;  
 PI WPI; 1995-302545/39.  
 DR Compn. inducing cytotoxic T lymphocyte response to pref. viral,  
 XX bacterial, parasitic or tumour antigens - useful in the treatment  
 PT and prevention of diseases associated with the antigen e.g.  
 PT hepatitis B  
 XX Disclosure; Page 17; 109pp; English.  
 PS A compsn. which induces a cytotoxic T lymphocyte (CTL) response to  
 CC an antigen (Ag) in a mammal comprises, a CTL Ag response inducing  
 CC peptide (i.e. R78824-R78853) and a lipid conjugated helper T cell  
 CC inducing peptide. The compsn. induces a CTL response to bacterial,  
 CC viral or tumour Ags, and is therefore useful in the treatment and  
 CC prevention of diseases associated with the Ag.  
 XX Sequence 8 AA;  
 SQ

Query Match 34.8%; Score 23; DB 16; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEEIDY 13  
 Db 3 kdeldy 8

## RESULT 14

W54663  
 ID W54663 standard; peptide; 8 AA.

XX AC W54663;  
 XX

DT 25-SEP-1998 (first entry)  
 XX

DE Peptide from P.falciparum CSP 368-375.  
 XX

KW Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;  
 KW vaccine; treatment.  
 XX

OS Synthetic.  
 XX

PN W09813378-A1.  
 XX

PD 02-APR-1998.  
 XX

PF 25-SEP-1997; 97WO-NL00536.  
 XX

PR 26-SEP-1996; 96EP-0202701.  
 XX

PA (UYLE-) RIJKSUNIV LEIDEN.  
 XX

PI Drijfhout JW, Koning F;  
 XX

DR WPI; 1998-230631/20.  
 XX

XX Increasing uptake and presentation of antigen(s) - by adding mannose  
 PT residue(s) to antigen for increasing T cell response, useful in,  
 PT e.g. vaccines against viral infection(s)  
 XX

PS Disclosure; Page 29; 47pp; English.  
 XX

CC The peptides W54559-W54809 are examples of peptides to which at least 1  
 CC (preferably 2) mannose can be attached to increase their uptake as  
 CC antigens by antigen-presenting cells. Uptake of agonist mannosylated  
 CC peptides will increase the T cell response, whereas uptake of antagonist  
 CC peptides blocks the T cell response. Blocking binding of immunogenic  
 CC autoantigens can be used in treatment of type I diabetes, rheumatoid  
 CC arthritis, graft rejection etc., also to induce T-cell non-  
 CC responsiveness. Vaccines containing mannosylated antigen are used to  
 CC prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths  
 CC and parasites.  
 XX

SQ Sequence 8 AA;

Query Match 34.8%; Score 23; DB 19; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEEIDY 13  
 Db 3 kdeldy 8

## RESULT 15

W54664

ID W54664 standard; peptide; 8 AA.

XX AC W54664;  
 XX

DT 25-SEP-1998 (first entry)  
 XX  
 DE Peptide from P.falciparum CSP 368-375 b.  
 XX

KW Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;  
 KW vaccine; treatment.  
 XX

OS Synthetic.  
 XX

PN W09813378-A1.  
 XX

PD 02-APR-1998.  
 XX

PF 25-SEP-1997; 97WO-NL00536.  
 XX

PR 26-SEP-1996; 96EP-0202701.  
 XX

PA (UYLE-) RIJKSUNIV LEIDEN.  
 XX

PI Drijfhout JW, Koning F;  
 XX

DR WPI; 1998-230631/20.  
 XX

XX Increasing uptake and presentation of antigen(s) - by adding mannose  
 PT residue(s) to antigen for increasing T cell response, useful in,  
 PT e.g. vaccines against viral infection(s)  
 XX

PS Disclosure; Page 29; 47pp; English.  
 XX

CC The peptides W54559-W54809 are examples of peptides to which at least 1  
 CC (preferably 2) mannose can be attached to increase their uptake as  
 CC antigens by antigen-presenting cells. Uptake of agonist mannosylated  
 CC peptides will increase the T cell response, whereas uptake of antagonist  
 CC peptides blocks the T cell response. Blocking binding of immunogenic  
 CC autoantigens can be used in treatment of type I diabetes, rheumatoid  
 CC arthritis, graft rejection etc., also to induce T-cell non-  
 CC responsiveness. Vaccines containing mannosylated antigen are used to  
 CC prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths  
 CC and parasites.  
 XX

SQ Sequence 8 AA;

Query Match 34.8%; Score 23; DB 19; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEEIDY 13  
 Db 3 kdeldy 8

Search completed: January 17, 2001, 13:34:23  
 Job time: 543 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:31:20 ; Search time 13.67 Seconds  
(without alignments)  
64,573 Million cell updates/sec

Title: US-08-765-837-1  
Perfect score: 66  
Sequence: 1 RTTQSDQEEIDY 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues  
Total number of hits satisfying chosen parameters: 1828

Minimum DB seq length: 0  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	30.3	13	2 T46862	hypothetical prote
2	20	30.3	13	2 S65612	tubulin alpha-chain
3	19	28.8	10	2 H61308	hemocyanin subunit
4	19	28.8	11	2 S19775	wound-induced prot
5	19	28.8	12	2 B56049	urinary tract ston
6	18	27.3	11	2 A40795	glycoprotein H-a -
7	18	27.3	11	2 S18385	NADP-cytochrome P4
8	18	27.3	12	2 S16335	beta-conglycinin a
9	18	27.3	12	2 S51737	T-cell receptor be
10	18	27.3	13	2 S47371	T-cell antigen rec
11	17	25.8	7	2 B39040	calsequestrin, fas
12	17	25.8	8	2 PA0035	protein QA300039 -
13	17	25.8	9	2 PH0108	late G1-69 protein
14	17	25.8	10	2 A61007	hementin (EC 3.4.-
15	17	25.8	10	2 S43625	cytochrome-c oxida
16	17	25.8	12	2 JQ2308	hypothetical 1.4K
17	17	25.8	12	2 JQ2318	hypothetical 1.4K
18	17	25.8	12	2 PH1605	Ig H chain V-D-J r
19	17	25.8	12	2 PH1611	Ig H chain V-D-J r
20	17	25.8	13	2 S15755	actin 7 - soybean
21	17	25.8	13	2 PC1149	equinatoxin 1A - s
22	17	25.8	13	2 S47373	T-cell antigen rec
23	17	25.8	13	2 I49637	deoxynucleotidyltr
24	16	24.2	7	2 A34818	vicilin 72K chain
25	16	24.2	9	2 PT0299	Ig heavy chain CRD
26	16	24.2	10	1 GMROL2	leucosulfakinin-II
27	16	24.2	10	2 S39392	calpain (EC 3.4.22
28	16	24.2	10	2 S48182	bacterioferritin -
29	16	24.2	10	2 B60656	leucosulfakinin II

30	16	24.2	12	2 S07206	kassinin - Senegal
31	16	24.2	12	2 PH1587	Ig H chain V-D-J r
32	16	24.2	13	2 S39413	tubulin beta chain
33	16	24.2	13	2 A61288	spore proteinase q
34	15	22.7	5	2 A32014	tram protein - Bac
35	15	22.7	6	2 B44510	hypothetical prote
36	15	22.7	6	2 B35640	cerebellar degen
37	15	22.7	7	2 S57274	triacylglycerol li
38	15	22.7	8	2 S63493	dissimilatory sulf
39	15	22.7	8	2 S21663	neuropeptide - flo
40	15	22.7	8	2 S65381	cytochrome-c oxida
41	15	22.7	10	2 PH0113	alpha-amylase (EC
42	15	22.7	10	2 PH0895	T-cell receptor be
43	15	22.7	11	2 S66606	quinoline 2-oxid
44	15	22.7	11	2 S65395	chemical-sense-rel
45	15	22.7	11	2 S43626	cytochrome-c oxida

ALIGNMENTS

RESULT 1  
T46862  
hypothetical protein algT [imported] - Pseudomonas aeruginosa (fragment)  
C:Species: Pseudomonas aeruginosa  
C:Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 21-Jul-2000  
C:Accession: T46862  
R:Devries, C.A.; Hassett, D.J.; Flynn, J.L.; Ohman, D.E.  
Gene 156, 63-67, 1995  
A:Title: Genetic linkage in Pseudomonas aeruginosa of algT and nadB: mutation in nadB  
A:Reference number: 224110; MUID:95255569  
A:Accession: T46862  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-13 <DEV>  
A:Cross-references: EMBL:U17232; NID:g608529; PIDN:AAA92355.1; PID:g608531  
A:Experimental source: strain FRD1; developmental stage mucoid (alginate overproducin  
C:Genetics:  
A>Note: algT

Query Match 30.3%; Score 20; DB 2; Length 13;  
Best Local Similarity 30.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 TQSDQEEID 12  
| : | : | :  
DB 3 TQEQDQLVE 12

RESULT 2  
S65612  
tubulin alpha-chain - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999  
C:Accession: S65612  
R:Ruediger, M.; Plessmann, U.; Ruediger, A.H.; Weber, K.  
FEBS Lett. 364, 147-151, 1995  
A:Title: Beta tubulin of bull sperm is polyglycylated.  
A:Reference number: S65611; MUID:95269788  
A:Accession: S65612  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-13 <RUE>

Query Match 30.3%; Score 20; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 DQEEID 12  
| : | : | :  
DB 1 DYEVE 6

```

RESULT 3
H61308
hemocyanin subunit IV - Atlantic horseshoe crab (fragment)
C:Species: Limulus polyphemus (Atlantic horseshoe crab)
C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
C:Accession: H61308
R:Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.
FEBS Lett. 106, 289-291, 1979
A:Title: Structural characterization of seven different subunits in Androctonus australis
A:Reference number: A61308; MUID:80047238
A:Accession: H61308
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <JOL>

Query Match 28.8%; Score 19; DB 2; Length 10;
Best Local Similarity 44.4%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLQSDQEEI 11
| | : | |
Db 1 TLKEKQDXI 9

RESULT 4
S19775
wound-induced protein - tomato (fragment)
C:Species: Lycopersicon esculentum (tomato)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Sep-1997
C:Accession: S19775
R:Parsons, B.L.
submitted to the EMBL Data Library, May 1991
A:Reference number: S19775
A:Accession: S19775
A:Molecule type: mRNA
A:Residues: 1-11 <PAR>
A:Cross-references: EMBL:X59884; NID:g19323; PID:g19324

Query Match 28.8%; Score 19; DB 2; Length 11;
Best Local Similarity 37.5%; Pred. No. 1.6e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 SDQEEIDY 13
| | : | |
Db 3 SSKKGLDY 10

RESULT 5
B56049
urinary tract stone matrix protein, 41.5K - unidentified organism (fragment)
C:Species: unidentified organism
C:Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 20-Apr-2000
C:Accession: B56049
R:Binette, J.P.; Binette, M.B.
Scanning Microsc. 8, 233-239, 1994
A:Title: Sequencing of proteins extracted from stones.
A:Reference number: A56049; MUID:95215817
A:Accession: B56049
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-12 <BIN>
A:Experimental source: urate-calcium oxalate kidney stones
A>Note: the source is designated as Homo sapiens, however the true source probably originated from

Query Match 28.8%; Score 19; DB 2; Length 12;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 TLQSDQEEI 11
| | : | |
Db 1 TLKEKQDXI 9

RESULT 6
A40795
glycoprotein H-a - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 31-Dec-1993
C:Accession: A40795
R:Christie, D.L.; Batchelor, D.C.; Palmer, D.J.
J. Biol. Chem. 266, 15679-15683, 1991
A:Title: Identification of kex2-related proteases in chromatin granules by partial a
A:Reference number: A40795; MUID:91340701
A:Accession: A40795
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-11 <CHR>
C:Keywords: glycoprotein

Query Match 27.3%; Score 18; DB 2; Length 11;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 DQBEID 12
| | : | |
Db 3 DINEID 8

RESULT 7
S18385
NADP-cytochrome P450 reductase-related protein - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Nov-1993 #sequence_revision 12-Apr-1996 #text_change 07-Feb-1997
C:Accession: S18385
R:Nadler, S.G.; Strobel, H.W.
Arch. Biochem. Biophys. 290, 277-284, 1991
A:Title: Identification and characterization of an NADPH-cytochrome P450 reductase de
A:Reference number: S18385; MUID:92027739
A:Accession: S18385
A:Molecule type: protein
A:Residues: 1-11 <NAD>
C:Keywords: NADP

Query Match 27.3%; Score 18; DB 2; Length 11;
Best Local Similarity 42.9%; Pred. No. 2.5e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 SDQEEID 12
| | : | |
Db 4 ADPEDVD 10

RESULT 8
S16335
beta-conglycinin alpha' chain - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 21-Nov-1993 #sequence_revision 27-Feb-1997 #text_change 13-Mar-1997
C:Accession: S16335
R:Coates, J.B.; Medeiros, J.S.; Thanh, V.H.; Nielsen, N.C.
Arch. Biochem. Biophys. 243, 184-194, 1985
A:Title: Characterization of the subunits of beta-conglycinin.
A:Reference number: S16334; MUID:86049421
A:Accession: S16335
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-12 <COA>

Query Match 27.3%; Score 18; DB 2; Length 12;

```

Best Local Similarity 22.2%; Pred. No. 2.7e+03;  
Matches 2; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 LQSDQEEID 12  
:::|::|:  
Db 1 VEEEEEEXE 9

## RESULT 9

T-cell receptor beta-chain joining region - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 05-Nov-1999  
C:Accession: S51737  
R:Durinovic-Bello, I.; Steinle, A.; Ziegler, A.G.; Schendel, D.J.  
submitted to the EMBL Data Library, November 1993  
A:Reference number: S51732  
A:Accession: S51737  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-12 <DUR>  
A:Cross-references: EMBL:Z28345; NID:g5607126; PIDN:CAA82199.1; PID:g5607127  
C:Keywords: T-cell receptor

Query Match 27.3%; Score 18; DB 2; Length 12;  
Best Local Similarity 37.5%; Pred. No. 2.7e+03;  
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 SDQEEIDY 13  
|::|:  
Db 4 SQEADIQY 11

## RESULT 10

T-cell antigen receptor VJ junction beta chain - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S47371  
R:Lehner, P.J.  
submitted to the EMBL Data Library, August 1994  
A:Description: Human HLA-A0201 restricted recognition of Influenza A is dominated by T  
A:Reference number: S47355  
A:Accession: S47371  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-13 <LEH>  
A:Cross-references: EMBL:Z35696; NID:g527483; PIDN:CAA84765.1; PID:g527484  
C:Keywords: T-cell receptor

Query Match 27.3%; Score 18; DB 2; Length 13;  
Best Local Similarity 37.5%; Pred. No. 3e+03;  
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTLSQDQE 9  
:::|::|:  
Db 3 SSIRSDGE 10

## RESULT 11

B39040  
calsequestrin, fast skeletal muscle - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 23-Feb-1997  
C:Accession: B39040  
R:Calaf, S.E.; Jones, L.R.  
J. Biol. Chem. 266, 391-398, 1991  
A:Title: Phosphorylation of cardiac and skeletal muscle calsequestrin isoforms by casein  
A:Reference number: A39040; MUID:91093153  
A:Accession: B39040  
A:Status: preliminary

A:Molecule type: protein  
A:Residues: 1-7 <CAL>  
C:Keywords: phosphoprotein; skeletal muscle

Query Match 25.8%; Score 17; DB 2; Length 7;  
Best Local Similarity 40.0%; Pred. No. 2e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 DQEEI 11  
|::|:  
Db 2 DEEDL 6

## RESULT 12

protein QA300039 - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 06-Jun-1997  
C:Accession: PA0035  
R:Kano, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.  
submitted to JIPID, July 1994  
A:Description: Separation and characterization of Arabidopsis proteins by two-dimensi  
A:Reference number: PA0001  
A:Accession: PA0035  
A:Molecule type: protein  
A:Residues: 1-8 <KAM>  
A:Experimental source: stem

Query Match 25.8%; Score 17; DB 2; Length 8;  
Best Local Similarity 37.5%; Pred. No. 2e+05;  
Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LQSDQEEI 11  
|::|:  
Db 1 LQXDNSXV 8

## RESULT 13

PH0108  
late G1-69 protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
C:Accession: PH0108  
R:Nikaido, T.; Bradley, D.W.; Pardee, A.B.  
Exp. Cell Res. 192, 102-109, 1991  
A:Title: Molecular cloning of transcripts that accumulate during the late G1 phase in  
A:Reference number: PH0108; MUID:91078351  
A:Accession: PH0108  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-9 <NIK>

Query Match 25.8%; Score 17; DB 2; Length 9;  
Best Local Similarity 33.3%; Pred. No. 2e+05;  
Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 QSDQEEIDY 13  
|::|:  
Db 1 EGDDEGEY 9

## RESULT 14

A61007  
hementin (EC 3.4.21.1) - Amazon leech (fragment)  
C:Species: Haementeria ghilianii (Amazon leech)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 07-May-1999  
C:Accession: A61007  
R:Swadesh, J.K.; Huang, I.Y.; Budzynski, A. Z.  
J. Chromatogr. 502, 359-369, 1990  
A:Title: Purification and characterization of hementin, a fibrinolytic protease t

A;Reference number: A61007; MUID:90256973  
A;Accession: A61007  
A;Molecule type: protein  
A;Residues: 1-10 <SWA>  
C;Keywords: anticoagulant; hydrolase; saliva

Query Match 25.8%; Score 17; DB 2; Length 10;  
Best Local Similarity 30.0%; Pred. No. 3.3e+03;  
Matches 3; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 TTLSQDQEEI 11  
||| : :  
Db 1 TTLTPEPEPL 10

## RESULT 15

S43625  
cytochrome-c oxidase (EC 1.9.3.1) chain Va, hepatic - rainbow trout (fragment)  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 20-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 18-Jul-1997  
C;Accession: S43625  
R;Freund, R.; Kadenbach, B.  
Eur. J. Biochem. 221, 1111-1116, 1994  
A;Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cytochrome-c oxidase  
A;Reference number: S43624; MUID:94237150  
A;Accession: S43625  
A;Molecule type: protein  
A;Residues: 1-10 <FRE>  
A;Note: the source is designated as Salmo gairdneri  
C;Genetics:  
A;Genome: nuclear  
C;Keywords: liver; membrane-associated complex; mitochondrion; oxidoreductase

Query Match 25.8%; Score 17; DB 2; Length 10;  
Best Local Similarity 33.3%; Pred. No. 3.3e+03;  
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 LOSDOE 9  
:::|  
Db 5 VETDEE 10

Search completed: January 17, 2001, 13:34:57  
Job time: 217 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:34:25 ; Search time 8.41 Seconds  
(without alignments)  
49.920 Million cell updates/sec

Title: US-08-765-837-1

Perfect score: 66

Sequence: 1 RTTLOSDQEEIDY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 518

Minimum DB seq length: 0

Maximum DB seq length: 13

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	19	28.8	10	1 SYK_CAMUP	Q46464 campylobact
2	18	27.3	12	1 UCRH_MOUSE	P99028 mus musculus
3	17	25.8	10	1 COXA_ONCMY	P80328 oncorhynchu
4	17	25.8	13	1 ACT7_SOYBN	P15987 glycine max
5	17	25.8	13	1 ORCK_ORCLI	P37086 orconectes
6	16	24.2	7	1 UF03_MOUSE	P38641 mus musculus
7	16	24.2	8	1 UPAL_HUMAN	P30087 homo sapien
8	16	24.2	10	1 LSK2_LEUMA	P09039 leucophaea
9	16	24.2	11	1 COXA_CANFA	P99501 canis famli
10	16	24.2	12	1 TKN_KASSE	P08611 kassina sen
11	15	22.7	5	1 TRM3_ECOLI	P13973 escherichia
12	15	22.7	7	1 UH1L_RAT	P56576 rattus norv
13	15	22.7	8	1 COXG_RAT	P80430 rattus norv
14	15	22.7	10	1 LCMS_LEUMA	P21144 leucophaea
15	15	22.7	10	1 SPI_HALRO	Q10997 halocynthia
16	15	22.7	11	1 Q2OA_COMTE	P80464 comamonas t
17	15	22.7	12	1 CALM_TETTH	Q05055 tetrahymena
18	15	22.7	13	1 LPAA_FORGI	P81411 porphyromon
19	14	21.2	10	1 FIBB_CERSI	P14537 ceratotheri
20	14	21.2	10	1 ODP2_BOVIN	P11180 bos taurus
21	14	21.2	10	1 UPA9_HUMAN	P30095 homo sapien
22	14	21.2	10	1 VEG6_BACSU	P80699 bacillus su
23	14	21.2	12	1 LOSK_LOCMI	P47733 locusta mig
24	13	19.7	8	1 CCN3_MACEU	P30369 macropus eu
25	13	19.7	9	1 FARL_CALVO	P41856 calliphora
26	13	19.7	9	1 FIBB_ERIPA	P19346 erythrocebu
27	13	19.7	9	1 FLA2_TREHY	P80159 treponema h
28	13	19.7	9	1 LPCA_STAAU	P36884 staphylococ
29	13	19.7	9	1 NSK1_SARBU	P41492 sarcophaga
30	13	19.7	9	1 UF02_MOUSE	P86640 mus musculus
31	13	19.7	10	1 CAER_LITXA	P56264 litorea xan
32	13	19.7	10	1 FARP_LOCMI	P38553 locusta mig
33	13	19.7	10	1 GON1_CHEPR	P80677 chelyosoma

34	13	19.7	10	1 NEMS_DROME	P41494 diosophila
35	13	19.7	10	1 UP11_CAEEL	P55956 caenorhabdi
36	13	19.7	11	1 HS70_PINPS	P81672 pinus pinus
37	13	19.7	11	1 LSK1_LEUMA	P04428 leucoplaea
38	13	19.7	11	1 LSKP_PERAM	P36885 periplaneta
39	13	19.7	12	1 UKA2_HUMAN	P31144 homo sapien
40	13	19.7	13	1 FIBB_RABIT	P14478 oryctolagus
41	13	19.7	13	1 PRC_SALTY	P43669 salmonealia
42	13	19.7	13	1 SODM_CANFA	P54712 canis famli
43	13	19.7	13	1 UN10_CLOPA	P81345 clostridium
44	12	18.2	8	1 AKH_MEML	P25423 melolontha
45	12	18.2	9	1 ULAH_HUMAN	P31934 homo sapien

## ALIGNMENTS

```
RESULT 1
SYK_CAMUP
ID SYK_CAMUP STANDARD; PRT; 10 AA.
AC Q46464;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSKS)
DE (FRAGMENT).
GN LYSS.
OS Campylobacter upsallensis.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43954;
RX MEDLINE=97149302; PubMed=8996110;
RA Bourke B., Rashid S.T., Bingham H.L., Chan V.L.;
RT "Characterization of Campylobacter upsallensis fur and its
RT localization in a highly conserved region of the Campylobacter
RT genome.";
RL Gene 183:219-224(1996).
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) -> AMP +
CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L77076; AAA4342.1; -.
CC INTERPRO: IPR002106; -.
DR PROSITE: P500179; AA_TRNA_LIGASE_IL1; PARTIAL.
DR PROSITE: P500339; AA_TRNA_LIGASE_IL2; PARTIAL.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1218 MW; 9CEA46AB13372B05 CRC64;
```

Query Match 28.8%; Score 19; DB 1; Length 10;  
Best Local Similarity 33.3%; Pred. No. 6.8e+02;  
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RTTLOSDQE 9  
:::|::|  
Db 2 KSELKEDKE 10

RESULT 2  
UCRH\_MOUSE

ID UCRH\_MOUSE STANDARD; PRT; 12 AA.  
AC P99028;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX II KDA PROTEIN (EC 1.10.2.2)  
DE (MITOCHONDRIAL HINGE PROTEIN) (CYTOCHROME C1, NONHEME 11 KDA PROTEIN)  
DE (COMPLEX III SUBUNIT VIII) (FRAGMENT).  
GN UOCHR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RC TISSUE=LIVER;  
RP Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,  
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,  
RA Cowthorne M.;  
RL Submitted (AUG-1998) to the SWISS-PROT data bank.  
CC -!- FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C  
CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH  
CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THIS PROTEIN MAY  
CC MEDIATE FORMATION OF THE COMPLEX BETWEEN CYTOCHROMES C AND C1.  
CC -!- CATALYTIC ACTIVITY: O<sub>2</sub>(2) + 2 FERRICYTOCHROME C = O +  
CC 2 FERROCYTOCHROME C.  
CC -!- SUBUNIT: B<sub>1</sub> COMPLEX CONTAINS 11 SUBUNITS; 3 RESPIRATORY SUBUNITS  
CC (CYTOCHROME B, CYTOCHROME C1, RIESKE PROTEIN), 2 CORE PROTEINS AND  
CC 6 LOW-MOLECULAR WEIGHT PROTEINS  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE UOCHR/OCR6 FAMILY.  
DR SWISS-2DPAGE; P99028; MOUSE.  
KW Mitochondrion; Inner membrane; Electron transport; Respiratory chain;  
KW Oxidoreductase.  
FT DOMAIN 5 10 POLY-GLU.  
FT NON\_TER 12 12  
FT SEQUENCE 12 AA; 1402 MW; 6B256F18801B1B1B CRC64;

Query Match 27.3%; Score 18; DB 1; Length 12;  
Best Local Similarity 28.6%; Pred. No. 1.2e+03;  
Matches 2; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 QSDQEEI 11  
: : : : :  
Db 5 EEEEEEL 11

RESULT 3  
COXA\_ONCMY  
ID COXA\_ONCMY STANDARD; PRT; 10 AA.  
AC P80328;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA (EC 1.9.3.1) (FRAGMENT).  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
RN [1]  
RP TISSUE=LIVER;  
RC MEDLINE=94237150; PubMed=8181469;  
RA Freund R., Kadenbach B.;  
RT Identification of tissue-specific isoforms for subunits Vb and VIIa  
RT of cytochrome c oxidase isolated from rainbow trout.";  
RL Eur. J. Biochem. 221:1111-1116(1994).  
CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C  
CC OXIDASE. THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.  
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
CC 4 FERRICYTOCHROME C.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.

DR PIR; S43625; S43625.  
KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1144 MW; C535C5B1AB02C33D CRC64;

Query Match 25.8%; Score 17; DB 1; Length 10;  
Best Local Similarity 33.3%; Pred. No. 1.5e+03;  
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 LQSDQE 9  
: : : : :  
Db 5 VETDEE 10

RESULT 4  
ACT7\_SOYBN  
ID ACT7\_SOYBN STANDARD; PRT; 13 AA.  
AC P15987;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE ACTIN 7 (FRAGMENT).  
GN SACT.  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; Cuscutids I;  
OC Fabales; Fabaceae; Papilionoideae; Glycine.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. WAYNE;  
RX MEDLINE=91346640; PubMed=2102831;  
RA Pearson L., Meagher R.B.;  
RT "Diverse soybean actin transcripts contain a large intron in the 5'  
RT untranslated leader: structural similarity to vertebrate muscle actin  
RT genes.";  
RL Plant Mol. Biol. 14:513-526(1990).  
CC -!- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED  
CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED  
CC IN ALL EUKARYOTIC CELLS.  
CC -!- FUNCTION: ESSENTIAL COMPONENT OF CELL CYTOSKELETON; PLAYS AN  
CC IMPORTANT ROLE IN CYTOPLASMIC STREAMING, CELL SHAPE DETERMINATION,  
CC CELL DIVISION, ORGANELLE MOVEMENT AND EXTENSION GROWTH.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- MISCELLANEOUS: THERE ARE AT LEAST 16 ACTIN GENES IN SOYBEAN.  
CC -!- SIMILARITY: BELONGS TO THE ACTIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X17120; CAA34980.1;  
DR PIR; S15755; S15755.  
DR INTERPRO; IPR000279;  
DR PROSITE; PS00406; ACTINS\_1; PARTIAL.  
DR PROSITE; PS00432; ACTINS\_2; PARTIAL.  
DR PROSITE; PS01132; ACTINS\_ACT LIKE; PARTIAL.  
KW Structural protein; Multigene family.  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1420 MW; 8BEFF3C36D4FD05A CRC64;

Query Match 25.8%; Score 17; DB 1; Length 13;  
Best Local Similarity 50.0%; Pred. No. 2e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 SQDEEI 11  
: : : : :  
Db 2 ADAEDI 7

RESULT 5  
ORCK\_ORCLI  
ID ORCK\_ORCLI STANDARD: PRT: 13 AA.  
AC P37086;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE ORCKININ.  
OS Orconectes limosus (Spinycheek crayfish).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;  
OC Astacoidae; Cambaridae; Orconectes.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=ABDOMINAL NERVE CORD;  
RX MEDLINE=93126144; PubMed=1480511;  
RA Stangler J., Hilbich C., Burdzik S., Keller R.;  
RT "Orckin: a novel myotropic peptide from the nervous system of the  
RT crayfish, Orconectes limosus.";  
RL Peptides 13:859-864(1992).  
CC -!- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND  
CC AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS.  
CC -!- TISSUE SPECIFICITY: ABDOMINAL NERVE CORD AND HINDGUT.  
KW Neuropeptide.  
SQ SEQUENCE 13 AA; 1518 MW; 8A318D7B4A93A40A CRC64;

Query Match 25.8%; Score 17; DB 1; Length 13;  
Best Local Similarity 75.0%; Pred. No. 2e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 EVID 12  
DB 3 DEID 6

RESULT 6  
UF03\_MOUSE  
ID UF03\_MOUSE STANDARD: PRT: 7 AA.  
AC P38641;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF FIBROBLASTS (P36) (FRAGMENT).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=FIBROBLAST;  
RX MEDLINE=95009907; PubMed=7523108;  
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;  
RT "Separation and sequencing of familial and novel murine proteins  
RT using preparative two-dimensional gel electrophoresis.";  
RL Electrophoresis 15:735-745(1994).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 5.1, ITS MW IS: 36 KDA.  
FT NON\_TER 7  
SQ SEQUENCE 7 AA; 842 MW; 6AA72B1DD81B1180 CRC64;

Query Match 24.2%; Score 16; DB 1; Length 7;  
Best Local Similarity 33.3%; Pred. No. 8.9e+04;  
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 DQEID 12  
DB 2 EEAELD 7

RESULT 7

UPAL\_HUMAN  
ID UPAL\_HUMAN STANDARD: PRT: 8 AA.  
AC P30087;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 2) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=PLASMA;  
RX MEDLINE=93092937; PubMed=1459097;  
RA Hughes G.J., Frutiger S., Pasquet N., Ravier F., Pasquall C.,  
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
RA Hochstrasser D.F.;  
RT "Plasma protein map: an update by microsequencing.";  
RL Electrophoresis 13:707-714(1992).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 4.9, ITS MW IS: 65 KDA.  
DR SWISS-2DPAGE; P30087; HUMAN.  
FT NON\_TER 1  
FT UNSURE 8  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 24.2%; Score 16; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 DOE 9  
DB 1 DOE 3

RESULT 8  
LSK2\_LEUMA  
ID LSK2\_LEUMA STANDARD: PRT: 10 AA.  
AC P09039;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE LEUCOSULFAKININ-II (LSK-II).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blaberoidea; Blaberidae; Leucophaea.  
RN [1]  
RP SEQUENCE.  
RC SPECIES=L.MADERAE;  
RX MEDLINE=87048769; PubMed=3778455;  
RA Nachman R.J., Holman G.M., Cook B.J., Haddon W.F., Ling N.;  
RT "Leucosulfaikin-II, a blocked sulfated insect neuropeptide with  
RT homology to cholecystokinin and gastrin.";  
RL Biochem. Biophys. Res. Commun. 140:357-364(1986).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=P.AMERICANA; TISSUE=CORPORA CARDIACA;  
RX MEDLINE=90137190; PubMed=2615921;  
RA Veenstra J.A.;  
RT "Isolation and structure of two gastrin/CCK-like neuropeptides from  
RT the American cockroach homologous to the leucosulfaikins.";  
RL Neuropeptides 14:145-149(1989).  
CC -!- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF  
CC THE COCKROACH HINDGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.  
CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.  
DR PIR; A26335; GMROL2.  
DR PIR; B60656; B60656.  
DR INTERPRO; IPR001651;  
DR PROSITE; PS00259; GASTRIN; 1.

KW Hormone; Amidation; Sulfatation.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 5 5 SULFATATION (IN L.MADERAE, BUT NOT IN  
 P.AMERICANA).  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1255 MW; 984F5391B86B5AAA CRC64;

Query Match 24.2%; Score 16; DB 1; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 2.3e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 5 QSDQEEIDY 13  
 : : : :  
 Db 1 QSD- - - -DY 5

RESULT 9  
 COXA\_CANFA  
 ID COXA\_CANFA STANDARD; PRT; 11 AA.  
 AC P99501;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE VA (EC 1.9.3.1) (FRAGMENT).  
 GN COX5A.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=HEART;  
 RX MEDLINE=9816340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
 dog heart proteins.";  
 RL Electrophoresis 18:2795-2802(1997).  
 CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C  
 OXIDASE. THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +  
 4 FERRICYTOCHROME C.  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.  
 DR HSC-2DPAGE; P99501; DOG.  
 KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.  
 FT NON\_TER 11 11  
 SEQUENCE 11 AA; 1274 MW; 910B35C5B1AB11F5 CRC64;

Query Match 24.2%; Score 16; DB 1; Length 11;  
 Best Local Similarity 40.0%; Pred. No. 2.5e+03;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 QSDQE 9  
 : : : :  
 Db 6 ETDEE 10

RESULT 10  
 TKN\_KASSE  
 ID TKN\_KASSE STANDARD; PRT; 12 AA.  
 AC P08611;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE KASSININ.  
 OS Kassina senegalensis (Senegal running frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Hyperoliidae;  
 OC Kassina.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77246385; PubMed=891753;

RA Anastasi A., Montecucchi P.C., Erspamer V., Visser J.;  
 RT "Amino acid composition and sequence of kassinin, a tachykinin  
 RT dodecapeptide from the skin of the African frog Kassina  
 RT senegalensis";  
 RL Experientia 33:857-858(1977).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR; S07206; S07206.  
 DR INTERPRO; IPR002040; -;  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.  
 FT MOD\_RES 12 12  
 SEQUENCE 12 AA; 1336 MW; 91757AB89DD6DAB5 CRC64;

Query Match 24.2%; Score 16; DB 1; Length 12;  
 Best Local Similarity 75.0%; Pred. No. 2.8e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 QSDQ 8  
 : : : :  
 Db 4 KSDQ 7

RESULT 11  
 TRM3\_ECOLI  
 ID TRM3\_ECOLI STANDARD; PRT; 5 AA.  
 AC P13973;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE TRAM PROTEIN (FRAGMENT).  
 GN TRAM.  
 OS Escherichia coli.  
 OG Plasmid IncFII R100.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88227859; PubMed=2836369;  
 RA Inamoto S., Yoshioka Y., Ohtsubo E.;  
 RT "Identification and characterization of the products from the traJ  
 RT and traI genes of plasmid R100.";  
 RL J. Bacteriol. 170:2749-2757(1988).  
 CC -!- FUNCTION: TRANSFER GENE PROTEIN. IS INVOLVED IN THE CONJUGATION  
 CC PROCESS OF BACTERIAL CELLS FOR THE EXCHANGE OF PLASMID DNA.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- SIMILARITY: TO TRAM PROTEIN OF OTHER PLASMIDS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M20941; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A32014; A32014.  
 KW Conjugation; Plasmid; DNA-binding.  
 FT NON\_TER 1 1  
 SEQUENCE 5 AA; 634 MW; 6B1B1AA443500000 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 5;  
 Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 QSDQE 9  
 : : : :  
 Db 1 1

Db 1 KNDEE 5

## RESULT 12

UHIL\_RAT  
 ID UHIL\_RAT STANDARD; PRT; 7 AA.  
 AC P56576;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF HEART TISSUE (SPOT P11) (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-WISTAR; TISSUE=HEART;  
 RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,  
 RA Jungblut P.R.;  
 RL Submitted (SEP-1998) to the SWISS-PROT data bank.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 8.5, ITS MW IS: 42 KDA.  
 FT UNSURE 2 2 OR A.  
 FT NON\_TER 7 7  
 SQ SEQUENCE 7 AA; 775 MW; 6866DB040DC5A6B0 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 7;

Best Local Similarity 50.0%; Pred. No. 8.9e+04; Indels 0; Gaps 0;

Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 QSDOE 10

ID 11:11:

Db 1 QSAREQ 6

## RESULT 13

COXG\_RAT  
 ID COXG\_RAT STANDARD; PRT; 8 AA.  
 AC P80430;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE VIB (EC 1.9.3.1) (AED) (FRAGMENT).  
 GN COX6B.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=WISTAR; TISSUE=LIVER;  
 RX MEDLINE=95324529; PubMed=7601105;  
 RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;  
 RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and  
 RT amino-terminal sequences suggest identity of the fetal heart and the  
 RT adult liver isoform."  
 RL Eur. J. Biochem. 230:235-241(1995).  
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE  
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN  
 CC MITOCHONDRIAL ELECTRON TRANSPORT.  
 CC -!- FUNCTION: THIS PROTEIN MAY BE ONE OF THE HEME-BINDING SUBUNITS OF  
 CC THE OXIDASE.  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC Oxidoreductase; Mitochondrion.  
 KW NON\_TER 1 1  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 1039 MW; 8101E9CA73AE456 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 8;

Best Local Similarity 33.3%; Pred. No. 8.9e+04; Indels 2; Gaps 0;

Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 QEEIDY 13

ID 11:11:

Db 1 QNXLDF 6

## RESULT 14

LCMS\_LEUMA  
 ID LCMS\_LEUMA STANDARD; PRT; 10 AA.  
 AC P21144; P41497;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE LEUCOMYOSUPPRESSIN (LMS) (LEM-MS).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidea; Blaberidae; Leucophaea.  
 RN [1]  
 RP SEQUENCE AND SYNTHESIS.  
 RC TISSUE=HEAD;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of leucomyosuppressin,  
 RT an insect neuropeptide that inhibits spontaneous contractions of the  
 RT cockroach hindgut."  
 RL Comp. Biochem. Physiol. 85C:329-333(1986).  
 CC -!- FUNCTION: INHIBITS THE SPONTANEOUS CONTRACTIONS OF COCKROACH  
 CC PROTODEUM (HINDGUT).  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1275 MW; D3C45229D2C1EAB2 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 10;

Best Local Similarity 20.0%; Pred. No. 3.4e+03; Indels 0; Gaps 0;

Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 9 EEIDY 13

ID 11:11:

Db 1 QDVODH 5

## RESULT 15

SPI\_HALRO  
 ID SPI\_HALRO STANDARD; PRT; 10 AA.  
 AC Q10997;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE SERINE PROTEINASE INHIBITOR (FRAGMENT).  
 OS Halocynthia roretzi (Sea squirt).  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 OC Stolidobranchia; Pyrosidae; Halocynthia.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=HEMOLYMPH;  
 RX MEDLINE=96321313; PubMed=8759295;  
 RA Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;  
 RT "Purification and characterization of a 58,000-Da proteinase  
 RT inhibitor from the hemolymph of a solitary ascidian, Halocynthia  
 RT roretzi."  
 RL Comp. Biochem. Physiol. 114B:1-9(1996).  
 CC -!- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 DR INTERPRO: IPR000215; ..  
 DR PROSITE: PS00284; SERPIN; PARTIAL.  
 KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1104 MW; 4225C73B1B187AA3 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.4e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 QSDQEE 10  
: | | |  
Db 2 KKDGE 7

Search completed: January 17, 2001, 13:35:49  
Job time: 84 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:34:25 ; Search time 19.04 Seconds

(without alignments)

80.026 Million cell updates/sec

Title: US-08-765-837-1

Perfect score: 66

Sequence: 1 RTTLQSDQEEIDY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 1632

Minimum DB seq length: 0

Maximum DB seq length: 13

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_15:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organella:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	33.3	13	5 Q24365	Q24365 drosophila
2	19	28.8	8	10 Q42507	Q42507 tritium ae
3	19	28.8	11	4 Q16427	Q16427 homo sapien
4	19	28.8	11	10 Q04131	Q04131 lycopersico
5	19	28.8	13	11 Q9QWU0	Q9QWU0 mus musculu
6	18	27.3	12	6 Q9TOY5	Q9TOY5 bos taurus
7	18	27.3	13	3 Q9URU0	Q9URU0 schizosacch
8	17	25.8	11	12 Q66874	Q66874 feline cali
9	17	25.8	11	12 Q66877	Q66877 feline cali
10	17	25.8	12	2 Q47251	Q47251 escherichia
11	17	25.8	12	12 Q78845	Q78845 human immun
12	17	25.8	13	5 Q9X242	Q9X242 albinaria h
13	17	25.8	13	5 Q9U5J3	Q9U5J3 trypanosoma
14	17	25.8	13	11 Q60517	Q60517 mus musculu
15	16	24.2	8	5 Q9N6M5	Q9N6M5 toxoplasma
16	16	24.2	9	2 Q9R735	Q9R735 streptomyce
17	16	24.2	9	13 Q9PS68	Q9PS68 gallus gall
18	16	24.2	10	2 Q9K343	Q9K343 escherichia
19	16	24.2	10	8 Q9XMB4	Q9XMB4 aegilops ta

20	16	24.2	10	11 Q9QVK7	Q9QVK7 mus sp. mep
21	16	24.2	11	2 Q47600	Q47600 escherichia
22	16	24.2	11	10 Q82070	Q82070 tritium ae
23	16	24.2	13	4 Q9UGG9	Q9UGG9 homo sapien
24	15	22.7	8	2 P72279	P72279 rhodococcus
25	15	22.7	8	4 Q9UK84	Q9UK84 homo sapien
26	15	22.7	9	4 P78484	P78484 homo sapien
27	15	22.7	9	8 Q9T2K9	Q9T2K9 spinacia oi
28	15	22.7	10	2 Q52762	Q52762 rhizobium s
29	15	22.7	10	5 P82384	P82384 drosophila
30	15	22.7	10	12 Q83978	Q83978 influenza a
31	15	22.7	11	2 Q9R5P3	Q9R5P3 serratia ma
32	15	22.7	11	4 Q9NQJ8	Q9NQJ8 homo sapien
33	15	22.7	12	2 Q02128	Q02128 desulfovibr
34	15	22.7	12	4 Q9UNV5	Q9UNV5 homo sapien
35	15	22.7	13	2 Q50038	Q50038 mycobacteri
36	15	22.7	13	2 Q47601	Q47601 escherichia
37	15	22.7	13	2 Q31295	Q31295 buchnera ap
38	15	22.7	13	2 Q9ZE21	Q9ZE21 buchnera ap
39	15	22.7	13	2 Q9R862	Q9R862 buchnera ap
40	15	22.7	13	4 Q9Y674	Q9Y674 homo sapien
41	15	22.7	13	11 Q63047	Q63047 rattus norv
42	15	22.7	13	13 P82097	P82097 litoria rub
43	15	22.7	13	13 P82098	P82098 litoria rub
44	14	21.2	8	2 Q9ZIE9	Q9ZIE9 neisseria m
45	14	21.2	8	5 Q9UB13	Q9UB13 albinaria h

## ALIGNMENTS

RESULT 1					
Q24365					
ID Q24365	PRELIMINARY:	PRT:	13 AA.		
AC Q24365					
DT 01-NOV-1996 (TReMBLrel. 01, Created)					
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)					
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)					
DE HYPOTHETICAL 1.5 KDA PROTEIN IN IMPL2 5'REGION.					
OS Drosophila melanogaster (Fruit fly).					
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					
OC Ephydroidea; Drosophilidae; Drosophila.					
OX NCBI_TaxID=7227;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=CANTON-S;					
RX MEDLINE=94139565; PubMed=8306886;					
RA Garbe J.C., Yang E., Fristrom J.W.;					
RT "IMP-L2: an essential secreted immunoglobulin family member implicated					
RL in neural and ectodermal development in Drosophila.";					
RL Development 119:1237-1250(1993).					
CC -!- CAUTION: THREE SHORT OPEN READING FRAMES ARE FOUND IN IMPL2 MRNA,					
CC LEADER REGION, THESE MAY BE TRANSLATED AS PART OF THE COMPLETE					
CC IMPL2 PROTEIN					
DR EMBL: L23066; AAB59249.1; .					
DR FLYBASE: FBgn0001257; Impl2.					
KW Hypothetical protein.					
SQ SEQUENCE 13 AA; 1486 MW; DE445043D063A736 CRC64;					

Query Match 33.3%; Score 22; DB 5; Length 13;  
Best Local Similarity 55.6%; Pred. No. 1.3e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; :ups 0;

QY	5	QSDQEEIDY 13
Db	3	QSSOLDIRY 11

RESULT 2					
Q42507					
ID Q42507	PRELIMINARY:	PRT:	8 AA.		

AC Q42507;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE HEAT SHOCK PROTEIN (FRAGMENT).  
 GN HSP70C OR HSP70A OR HSP70B.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV MUSTANG;  
 RX MEDLINE=96189275; PubMed=8605312;  
 RA Joshi C.P., Kumar S., Nguyen H.T.;  
 RT "Application of modified differential display technique for cloning  
 RT and sequencing of the 3' region from three putative members of wheat  
 RT HSP70 gene family."  
 RL Plant Mol. Biol. 30:641-646(1996).  
 DR EMBL; L41507; AAB02333.1; -;  
 DR EMBL; L41505; AAB02331.1; -;  
 DR EMBL; L41506; AAB02332.1; -;  
 KW Heat shock.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 886 MW; 71B2CB1B10532768 CRC64;

Query Match 28.8%; Score 19; DB 10; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 9 EVID 12  
 Db 5 EVID 8

RESULT 3  
 ID Q16427 PRELIMINARY; PRT; 11 AA.  
 AC Q16427;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE DYSTROPHIN PROTEIN (FRAGMENT).  
 GN DYSTROPHIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96163501; PubMed=8566960;  
 RA Holder E., Maeda M., Bies R.D.;  
 RT "Expression and regulation of the dystrophin Purkinje promoter in  
 RT human skeletal muscle, heart, and brain."  
 RL Hum. Genet. 97:232-239(1996).  
 DR EMBL; S81419; AAD14362.1; -;  
 FT NON\_TER 11  
 SQ SEQUENCE 11 AA; 1299 MW; DDC84321AB5A5A2 CRC64;

Query Match 28.8%; Score 19; DB 4; Length 11;  
 Best Local Similarity 37.5%; Pred. No. 3.7e+03;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LQSDQEEI 11  
 Db 4 VSSDEREM 11

RESULT 4  
 Q04131 PRELIMINARY; PRT; 11 AA.  
 ID Q04131

AC Q04131;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE WOUND-INDUCED PROTEIN (FRAGMENT).  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PIK-RED; TISSUE=PERICARP;  
 RX MEDLINE=91355936; PubMed=1715787;  
 RA Parsons B.L., Mattoo A.K.;  
 RT "Wound-regulated accumulation of specific transcripts in tomato fruit:  
 RT interactions with fruit development, ethylene and light."  
 RL Plant Mol. Biol. 17:453-464(1991).  
 DR EMBL; X59884; CAA42539.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 11 AA; 1278 MW; 92CB257828733325 CRC64;

Query Match 28.8%; Score 19; DB 10; Length 11;  
 Best Local Similarity 37.5%; Pred. No. 3.7e+03;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 SDQEEIDY 13  
 Db 3 SSKKGLDY 10

RESULT 5  
 ID Q9QWU0 PRELIMINARY; PRT; 13 AA.  
 AC Q9QWU0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE UNKNOWN (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129;  
 RA Rowen L., Qin S., Madan A., Loretz C., Hall J., James R., Dors M.,  
 RA Shaffer T., Abbasi N., Ratcliffe A., Dickhoff R., Lasky S., Hood L.;  
 RT "Sequence of the mouse major histocompatibility complex class II  
 RT region."  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF110520; AAC97964.1; -;  
 FT NON\_TER 13  
 SQ SEQUENCE 13 AA; 1653 MW; 832A1D689E825B1A CRC64;

Query Match 28.8%; Score 19; DB 11; Length 13;  
 Best Local Similarity 60.0%; Pred. No. 4.4e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 9 EEIDY 13  
 Db 6 DEYDY 10

RESULT 6  
 ID Q9TOY5 PRELIMINARY; PRT; 12 AA.  
 AC Q9TOY5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)



DE GLYCOPROTEIN H-A N-TERMINAL, GPH-A N-TERMINAL=KEX2/SUBTILISIN-RELATED  
DE PROTEASE.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RN SEQUENCE.  
RP MEDLINE=91340701; PubMed=1874725;  
RX Christie D.L., Batchelor D.C., Palmer D.J.;  
RA "Identification of kex2-related proteases in chromaffin granules by  
RT partial amino acid sequence analysis."  
RL J. Biol. Chem. 266:15679-15683(1991).  
SQ SEQUENCE 12 AA; 1374 MW; 0BDF36703B5B1440 CRC64;

Query Match 27.3%; Score 18; DB 6; Length 12;  
Best Local Similarity 66.7%; Pred. No. 6.1e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 DQEEID 12  
Db 4 DINEID 9

RESULT 7  
Q9URU0 PRELIMINARY; PRT; 13 AA.  
AC Q9URU0;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DE 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
DE PUTATIVE MITOCHONDRIAL IMPORT RECEPTOR SUBUNIT (FRAGMENT).  
GN SPC27B12.13.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;  
OC Schizosaccharomycetaceae; Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=972H-;  
RA Lyne M., Rajandream M.A., Barrell B.G., Lauber J., Hilbert H.,  
RA Duesterhoeft A;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL022072; CAB64681.1; -  
KW Receptor.  
FT NON\_TER 1  
SQ SEQUENCE 13 AA; 1456 MW; DCD0A7CCD2FA36C6 CRC64;

Query Match 27.3%; Score 18; DB 3; Length 13;  
Best Local Similarity 25.0%; Pred. No. 6.6e+03;  
Matches 2; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 LQSDQEEI 11  
Db 1 IQQQQQQL 8

RESULT 8  
Q66874 PRELIMINARY; PRT; 11 AA.  
AC Q66874;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DE 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE POLYMERASE (FRAGMENT).  
OS Feline calicivirus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Calicivirus.  
OX NCBI\_TaxID=11978;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=255;  
RX MEDLINE=95250311; PubMed=7732664;  
RA Seal B.S., Neill J.D.;  
RT "Capsid protein gene sequence of feline calicivirus isolates 255 and  
RT LLK: further evidence for capsid protein configuration among feline  
RT caliciviruses."  
RL Virus Genes 9:183-187(1995).  
DR EMBL; U07130; AAA74412.1; -  
FT NON\_TER 1  
SQ SEQUENCE 11 AA; 1232 MW; 4E8E5F5C0AAEB1E7 CRC64;

Query Match 25.8%; Score 17; DB 12; Length 11;  
Best Local Similarity 40.0%; Pred. No. 8.5e+03;  
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 TTLQSDQEEI 11  
Db 2 TALHCDVFEV 11

RESULT 9  
Q66877 PRELIMINARY; PRT; 11 AA.  
AC Q66877;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DE 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE POLYMERASE (FRAGMENT).  
OS Feline calicivirus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Calicivirus.  
OX NCBI\_TaxID=11978;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=LLK;  
RX MEDLINE=95250311; PubMed=7732664;  
RA Seal B.S., Neill J.D.;  
RT "Capsid protein gene sequence of feline calicivirus isolates 255 and  
RT LLK: further evidence for capsid protein configuration among feline  
RT caliciviruses."  
RL Virus Genes 9:183-187(1995).  
DR EMBL; U07131; AAA74415.1; -  
FT NON\_TER 1  
SQ SEQUENCE 11 AA; 1264 MW; 5E9F43EBCAAEB1E7 CRC64;

Query Match 25.8%; Score 17; DB 12; Length 11;  
Best Local Similarity 44.4%; Pred. No. 8.5e+03;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 TLQSDQEEI 11  
Db 3 TLHCDVFEV 11

RESULT 10  
Q47251 PRELIMINARY; PRT; 12 AA.  
AC Q47251;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
DE HEMC PROTEIN (FRAGMENT).  
GN HEMC.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=K12;

```
RX MEDLINE=88096587; PubMed=3320969;
RA Jordan P.M., Mgeje B.I.A., Alwan A.F., Thomas S.D.;
RT "Nucleotide sequence of hemD, the second gene in the hem operon of
RL Escherichia coli K-12";
RL Nucleic Acids Res. 15:10583-10583(1987).
DR EMBL; Y00883; CAA68775.1; -.
FT NON_TER 1
FT NON_TER 12
FT NON_TER 13
SQ SEQUENCE 12 AA; 1375 MW; 5D2DE8339BA045B3 CRC64;

Query Match 25.8%; Score 17; DB 2; Length 12;
Best Local Similarity 44.4%; Pred. No. 9.2e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RTTLQSDQE 9
   |::|::|
Db 1 RSSLSKSITE 9

RESULT 11
Q78845 PRELIMINARY; PRT; 12 AA.
AC Q78845;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE REV PROTEIN (FRAGMENT).
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Wong-Staal F., Gallo R.C., Chang N.T., Ghrayeb J., Papas T.S.,
RA Lautenberger J.A., Pearson M.L., Petteway S.R. Jr., Ivanoff L.,
RA Baumeister K., Whitehorn E.A., Rafalski J.A., Doran E.R.,
RA Josephs S.J., Starcich B., Livak K.J., Patarca R., Haseltine W.A.,
RA Ratner L.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III."
RL Nature 313:277-284(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA van Beveren C.P., Coffin J., Hughes S.;
RL (In) Weiss R.L., Teich N., Varmus H., Coffin J. (eds.);
RL RNA tumor viruses, second edition, 2, pp.2:1102-1123,
RL Cold Spring Harbor Laboratory Press, New York (1985).
DR EMBL; K02012; AAA44657.1; -.
FT NON_TER 12
FT NON_TER 13
SQ SEQUENCE 12 AA; 1266 MW; 8F6D0408CBAA865 CRC64;

Query Match 25.8%; Score 17; DB 12; Length 12;
Best Local Similarity 75.0%; Pred. No. 9.2e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 SDOE 9
   |::|
Db 8 SDEE 11

RESULT 12
Q9XZ42 PRELIMINARY; PRT; 13 AA.
AC Q9XZ42;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CALMODULIN (FRAGMENT).
OS Albinaria hippolyti.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Stylomatophora;
OC Clausiliidae; Alopinae; Albinaria.
OX NCBI_TaxID=69418;
```

```
RN SEQUENCE FROM N.A.
RA Schilthuis M., Hoekstra R.F., Gittenberger E.;
RT "Selective maintenance of a rare Cam haplotype in a land snail hybrid
RT zone.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF132317; AAD34387.1; -.
DR HSSP; P02593; IAJI.
FT NON_TER 1
FT NON_TER 13
FT NON_TER 13
SQ SEQUENCE 13 AA; 1463 MW; A7BDD19B8A114056 CRC64;

Query Match 25.8%; Score 17; DB 5; Length 13;
Best Local Similarity 44.4%; Pred. No. 1e+04;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LQSDQEEID 12
   |::|
Db 3 LQDMINEVD 11

RESULT 13
Q9U5J3 PRELIMINARY; PRT; 13 AA.
AC Q9U5J3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HEAT SHOCK PROTEIN 100 (FRAGMENT).
OS HSP100.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RA Webb H.D., Gaud A.F., Carrington M.;
RT "The trypanosoma brucei GPI-PLC gene is not linked to functionally
RT related genes or to genes showing the same developmentally regulated
RT expression.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250727; CAB60092.1; -.
KW Heat shock.
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1524 MW; 37C19EC3787FA728 CRC64;

Query Match 25.8%; Score 17; DB 5; Length 13;
Best Local Similarity 57.1%; Pred. No. 1e+04;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 LQSDQEE 10
   |::|
Db 7 LTSDEWE 13

RESULT 14
Q60517 PRELIMINARY; PRT; 13 AA.
AC Q60517;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 03, Last annotation update)
DE DEOXYNUCLEOTIDYLTRANSFERASE (FRAGMENT).
OS DNTT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87213162; PubMed=3579900;
```

RA Koikwai O., Kaneda T., Morishita R.;  
RT "Analysis of human terminal deoxynucleotidyl transferase cDNA  
expressible in mammalian cells."; 144:185-190(1987).  
RL Biochem. Biophys. Res. Commun. 144:185-190(1987).  
DR EMBL; M26145; AAA74592.1; -.  
KW Transferase.  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1508 MW; 53F7C770F617BAA7 CRC64;

Query Match 25.8%; Score 17; DB 11; Length 13;  
Best Local Similarity 60.0%; Pred. No. 1e+04;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEID 12  
Db 3 QELD 7

RESULT 15  
O9N6M5 PRELIMINARY; PRT: 8 AA.  
AC O9N6M5;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE DIHYDROFOLATE REDUCTASE THYMIDYLATE SYNTHASE (FRAGMENT).  
GN FOLI.  
OS Toxoplasma gondii.  
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;  
OC Toxoplasma.  
OX NCBI\_TaxID=5811;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RH, COUGAR TC751G34, SEA OTTER TC828G1, AND BEVERLEY;  
RA Lehmann T., Blackston C.R., Parmley S.F., Remington J.S., Dubey J.P.;  
RT "Strain Typing of Toxoplasma gondii: Comparison of Antigen-Coding and  
Housekeeping Genes";  
RL J. Parasitol. 0:0-0(2000).  
DR EMBL; AF249695; AAF79153.1; -.  
DR EMBL; AF249692; AAF79150.1; -.  
DR EMBL; AF249693; AAF79151.1; -.  
DR EMBL; AF249694; AAF79152.1; -.  
FT NON\_TER 1  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 1035 MW; 33CAAAA05B133044 CRC64;

Query Match 24.2%; Score 16; DB 5; Length 8;  
Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 EEID 12  
Db 3 KEID 6

Search completed: January 17, 2001, 13:36:13  
Job time: 108 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:26:41 ; Search time 12.27 Seconds  
(without alignments)  
19.025 Million cell updates/sec

Title: US-08-765-837-1

Perfect score: 66

Sequence: 1 RTTLOSQEEIDY 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 66855

Minimum DB seq length: 0

Maximum DB seq length: 13

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA: \*  
1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep: \*  
2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep: \*  
3: /cgn2.6/ptodata/2/iaa/6\_COMB.pep: \*  
4: /cgn2.6/ptodata/2/iaa/pctus\_COMB.pep: \*  
5: /cgn2.6/ptodata/2/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	34	51.5	11	3	US-08-441-935-28
2	30	45.5	11	3	US-08-441-935-22
3	25	37.9	12	5	5395759-7
4	25	37.9	13	2	US-08-140-137A-17
5	24	36.4	10	1	US-08-119-361-3
6	24	36.4	10	2	US-08-518-967-7
7	24	36.4	11	1	US-08-197-793-9
8	24	36.4	11	2	US-08-636-176-9
9	24	36.4	11	4	PCT-US95-01618-9
10	23	34.8	8	2	US-08-318-856A-3
11	23	34.8	8	2	US-08-318-856A-4
12	23	34.8	8	4	PCT-US95-02121-17
13	23	34.8	9	1	US-07-779-049-1
14	23	34.8	9	1	US-08-080-240-1
15	23	34.8	9	2	US-08-318-856A-20
16	23	34.8	9	2	US-08-318-856A-27
17	23	34.8	10	2	US-08-140-137A-4
18	23	34.8	12	1	US-08-367-758B-3
19	23	34.8	12	2	US-08-909-735-3
20	23	34.8	13	1	US-08-052-681-11
21	22	33.3	6	1	US-08-704-170-116
22	22	33.3	6	4	PCT-US94-02631-116
23	22	33.3	7	4	PCT-US93-08214-15
24	22	33.3	9	2	US-08-986-234-40
25	22	33.3	10	1	US-08-318-970B-8
26	22	33.3	10	2	US-08-140-137A-3
27	22	33.3	10	2	US-08-603-753D-9
28	22	33.3	10	2	US-08-603-753D-10

29 22 33.3 10 3 US-09-099-753-9 Sequence 9, Appli  
30 22 33.3 10 3 US-09-099-753-10 Sequence 10, Appl  
31 22 33.3 11 3 US-09-273-565-50 Sequence 50, Appl  
32 22 33.3 12 2 US-08-553-257A-24 Sequence 24, Appl  
33 21 31.8 7 1 US-08-704-170-18 Sequence 18, Appl  
34 21 31.8 7 3 US-09-040-216-39 Sequence 39, Appl  
35 21 31.8 7 4 PCT-US94-02631-18 Sequence 18, Appl  
36 21 31.8 8 5 5210075-13 Patent No. 5210075  
37 21 31.8 9 5 5210075-12 Patent No. 5210075  
38 21 31.8 10 2 US-08-139-609-8 Sequence 8, Appli  
39 21 31.8 10 5 5210075-11 Patent No. 5210075  
40 21 31.8 11 2 US-08-162-081B-3 Sequence 3, Appli  
41 21 31.8 11 2 US-08-235-515A-6 Sequence 6, Appli  
42 21 31.8 11 2 US-08-780-872-3 Sequence 3, Appli  
43 21 31.8 12 1 US-07-909-122-8 Sequence 8, Appli  
44 21 31.8 12 1 US-08-325-071-18 Sequence 18, Appli  
45 21 31.8 12 3 US-08-742-243-25 Sequence 25, Appli

#### ALIGNMENTS

RESULT 1  
US-08-441-935-28  
; Sequence 28, Application US/08441935  
; Patent No. 6060447  
; GENERAL INFORMATION:  
; APPLICANT: CHAPMAN, BARBARA  
; APPLICANT: BURKE, RAE LYN  
; APPLICANT: RASMUSSEN, MIRELLA EZBAN  
; APPLICANT: MIKKELSON, JAN MOLLER  
; TITLE OF INVENTION: PROTEIN COMPLEXES HAVING FACTOR VIII:C  
; TITLE OF INVENTION: ACTIVITY AND PRODUCTION THEREOF  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: REED & ROBINS  
; STREET: 285 HAMILTON AVENUE, SUITE 200  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,935  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/266,170  
; FILING DATE: 27-JUNE-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/652,099  
; FILING DATE: 7-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/051,916  
; FILING DATE: 19-MAY-1987  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/822,989  
; FILING DATE: 27-JAN-1986  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 2300-0048.11  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 327-3400  
; TELEFAX: (415) 327-3231  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-441-935-28

Query Match 51.5%; Score 34; DB 3; Length 11;  
Best Local Similarity 87.5%; Pred. No. 2.2;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RTTLOSDQ 8  
Db 4 RTXLQSDQ 11  
RESULT 2  
US-08-441-935-22  
Sequence 22, Application US/08441935  
Patent No. 6060447  
GENERAL INFORMATION:  
APPLICANT: CHAPMAN, BARBARA  
APPLICANT: BURKE, RAE LYN  
APPLICANT: RASMUSSEN, MIRELLA EZBAN  
APPLICANT: MIRKELSON, JAN MOLLER  
TITLE OF INVENTION: PROTEIN COMPLEXES HAVING FACTOR VIII:C  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 285 HAMILTON AVENUE, SUITE 200  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/441,935  
FILING DATE: 16-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 08/266,170  
FILING DATE: 27-JUNE-1994  
PRIOR APPLICATION DATA: US 07/652,099  
FILING DATE: 7-FEB-1991  
APPLICATION NUMBER: US 07/051,916  
FILING DATE: 19-MAY-1987  
PRIOR APPLICATION DATA: US 06/822,989  
FILING DATE: 27-JAN-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 2300-0048.11  
TELEPHONE: (415) 327-3400  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-441-935-22

Query Match 45.5%; Score 30; DB 3; Length 11;

Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RTTLOSDQ 8  
Db 4 RTXLQSDQ 11

RESULT 3  
5395759-7  
Patent No. 5395759  
APPLICANT: HOLMES, IAN H.; DYALL-SMITH, MICHAEL L.  
TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCE  
ENCODING THE HUMAN ROTAVIRUS MAJOR OUTER CAPSID GLYCOPROTEIN  
NUMBER OF SEQUENCES: 14  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/474,642  
FILING DATE: 29-APR-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 824,704  
FILING DATE: 04-FEB-1987  
SEQ ID NO: 7;  
LENGTH: 12  
5395759-7

Query Match 37.9%; Score 25; DB 5; Length 12;  
Best Local Similarity 55.6%; Pred. No. 93;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLQSDQEEI 11  
Db 4 TLQDMSSEL 12

RESULT 4  
US-08-140-137A-17  
Sequence 17, Application US/08140137A  
Patent No. 5817617  
GENERAL INFORMATION:  
APPLICANT: TUOMANEN, ELAINE  
APPLICANT: MASURE, H. R.  
TITLE OF INVENTION: ANALOGS OF ENDOTHELIAL LEUKOCYTE  
ADHESION MOLECULE (ELAM)  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/140,137A  
FILING DATE: 27-MAY-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-096  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid

;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; FRAGMENT TYPE: internal  
US-08-140-137A-17

Query Match 37.9%; Score 25; DB 2; Length 13;  
Best Local Similarity 41.7%; Pred. No. 1e+02;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 TLLQSDQEEIDY 13  
| : |||:  
Db 1 TLVAIQNEEIEY 12

RESULT 5  
US-08-119-361-3  
; Sequence 3, Application US/08119361  
; Patent No. 5523390  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; TITLE OF INVENTION: Porphyromonas gingivalis Arginine-specific Proteinase  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,361  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33878  
REFERENCE/DOCKET NUMBER: 21-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

US-08-119-361-3  
Query Match 36.4%; Score 24; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLLQSDQEEI 11  
||| |||  
Db 1 TLQKKIEEI 9

RESULT 6  
US-08-119-361-3  
; Sequence 3, Application US/08119361  
; Patent No. 5523390  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; TITLE OF INVENTION: Porphyromonas gingivalis Arginine-specific Proteinase  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,361  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33878  
REFERENCE/DOCKET NUMBER: 21-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

US-08-119-361-3  
Query Match 36.4%; Score 24; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLLQSDQEEI 11  
||| |||  
Db 1 TLQKKIEEI 9

RESULT 6  
US-08-119-361-3  
; Sequence 3, Application US/08119361  
; Patent No. 5523390  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; TITLE OF INVENTION: Porphyromonas gingivalis Arginine-specific Proteinase  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,361  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33878  
REFERENCE/DOCKET NUMBER: 21-93  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303-499-8080  
TELEFAX: 303-499-8089  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

US-08-119-361-3  
Query Match 36.4%; Score 24; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLLQSDQEEI 11  
||| |||  
Db 1 TLQKKIEEI 9

RESULT 6  
US-08-119-361-3  
; Sequence 3, Application US/08119361  
; Patent No. 5523390  
; GENERAL INFORMATION:  
; APPLICANT: Travis, James  
; APPLICANT: Potempa, Jan  
; APPLICANT: Barr, Philip J.  
; APPLICANT: Pavloff, Nadine  
; TITLE OF INVENTION: Porphyromonas gingivalis Arginine-specific Proteinase  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee and Winner, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80303

US-08-518-967-7  
; Sequence 7, Application US/08518967  
; Patent No. 5861307  
; GENERAL INFORMATION:  
; APPLICANT: Kuchino, Y. et al.  
; TITLE OF INVENTION: HUMAN s-myc-LIKE POLYPEPTIDE AND  
; TITLE OF INVENTION: A GENE CODING FOR SAID POLYPEPTIDE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIELDS & LEMACK  
; STREET: 176 East Main Street, Suite 8  
; CITY: Westboro  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 01581  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/518,967  
FILING DATE: August 24, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-207236  
FILING DATE: 31-AUG-94  
ATTORNEY/AGENT INFORMATION:  
NAME: Lemack, Kevin S.  
REGISTRATION NUMBER: 32,579  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 898-1818  
TELEFAX: (508) 898-2020  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-518-967-7

Query Match 36.4%; Score 24; DB 2; Length 10;  
Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 SDOEEID 12  
| : |||:  
Db 3 SEGEED 9

RESULT 7  
US-08-197-793-9  
; Sequence 9, Application US/08197793  
; Patent No. 5510461  
; GENERAL INFORMATION:  
; APPLICANT: Meuer, S.  
; APPLICANT: Schraven, B.  
; APPLICANT: Schoenhaut, D.  
; APPLICANT: Ratnofsky, S.  
; TITLE OF INVENTION: pp32: A Newly Identified CD45-Associated  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,793
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,019;
; FILING DATE: 19-APR-1991
; APPLICATION NUMBER: 08/004,199
; FILING DATE: 13-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-006CNCPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-197-793-9

Query Match 36.4%; Score 24; DB 1; Length 11;
Best Local Similarity 42.9%; Pred. No. 1.3e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 DQEEIDY 13
DB 2 DEQTDY 8

RESULT 8
US-08-636-176-9
; Sequence 9, Application US/08636176
; Patent No. 5848622
; GENERAL INFORMATION:
; APPLICANT: Meuer, S.
; APPLICANT: Schraven, B.
; APPLICANT: Schoenhaut, D.
; APPLICANT: Ratofsky, S.
; TITLE OF INVENTION: pp32: A Newly Identified CD45-Associated
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/636,176
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,019; 08/004,199
; FILING DATE: 19-APR-1991; 13-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-006CNCPC
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-636-176-9

Query Match 36.4%; Score 24; DB 2; Length 11;
Best Local Similarity 42.9%; Pred. No. 1.3e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 DQEEIDY 13
DB 2 DEQTDY 8

RESULT 9
PCT-US95-01618-9
; Sequence 9, Application PC/TUS9501618
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: pp32: A Newly Identified CD45-Associated
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01618
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US95 08/197,793
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-006CPCPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; PCT-US95-01618-9

Query Match 36.4%; Score 24; DB 4; Length 11;
Best Local Similarity 42.9%; Pred. No. 1.3e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 DQEEIDY 13
DB 2 DEQTDY 8
```



```
RESULT 10
US-08-318-856A-3
; Sequence 3, Application US/08318856A
; Patent No. 5972351
; GENERAL INFORMATION:
; APPLICANT: Adrian V.S. Hill, et al.
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,856A
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 263-PP1R15770US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 721-8200
; TELEFAX: (202) 721-8250
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-08-318-856A-3

Query Match 34.8%; Score 23; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEEIDY 13
Db 3 KDELDY 8

RESULT 11
US-08-318-856A-4
; Sequence 4, Application US/08318856A
; Patent No. 5972351
; GENERAL INFORMATION:
; APPLICANT: Adrian V.S. Hill, et al.
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)
; NUMBER OF SEQUENCES: 86
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,856A
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 263-PP1R15770US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 721-8200
; TELEFAX: (202) 721-8250
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-08-318-856A-4

Query Match 34.8%; Score 23; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEEIDY 13
Db 3 KDELDY 8

RESULT 12
PCT-US95-02121-17
; Sequence 17, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA: US 07/874,491  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992  
PRIOR APPLICATION DATA: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 457-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
PCT-US95-02121-17

Query Match 34.8%; Score 23; DB 4; Length 8;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEEIDY 13  
DB 3 KDELIDY 8

RESULT 13  
US-07-779-049-1  
Sequence 1, Application US/07779049  
Patent No. 5310659  
GENERAL INFORMATION:  
APPLICANT: ARAMORI, ICHIRO  
APPLICANT: FURAGAWA, MASAO  
APPLICANT: ONO, HIROKI  
APPLICANT: ISHITANI, YOSUKE  
APPLICANT: TSUMURA, MANA  
APPLICANT: IWAMI, MORITA  
APPLICANT: KOJO, HITOSHI  
TITLE OF INVENTION: NOVEL GL-7ACA ACYLASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/779,049  
FILING DATE: 19911018  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5310659man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-723-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500

TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-07-779-049-1

Query Match 34.8%; Score 23; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 QSDQEEID 12  
DB 1 QSDQEKAE 8

RESULT 14  
US-08-080-240-1  
Sequence 1, Application US/08080240  
Patent No. 5312750  
GENERAL INFORMATION:  
APPLICANT: ARAMORI, ICHIRO  
APPLICANT: FURAGAWA, MASAO  
APPLICANT: ONO, HIROKI  
APPLICANT: ISHITANI, YOSUKE  
APPLICANT: TSUMURA, MANA  
APPLICANT: IWAMI, MORITA  
APPLICANT: KOJO, HITOSHI  
TITLE OF INVENTION: NOVEL GL-7ACA ACYLASE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/080,240  
FILING DATE: 19930622  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5312750man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-803-0 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)413-3000  
TELEFAX: (703)413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-080-240-1

Query Match 34.8%; Score 23; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 QSDQEID 12  
Db 1 QSEQEKAE 8

## RESULT 15

US-08-318-856A-20  
; Sequence 20, Application US/08318856A  
; Patent No. 5972351  
; GENERAL INFORMATION:  
; APPLICANT: Adrian V.S. Hill, et al.  
; TITLE OF INVENTION: PLASMIDIUM FALCIPARUM MHC CLASS I-  
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE  
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1+  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,856A  
; FILING DATE: October 3, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 92 08 068.8  
; FILING DATE: April 3, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 92 17 704.7  
; FILING DATE: August 20, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/GB93/00711  
; FILING DATE: April 5, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee Cheng  
; REGISTRATION NUMBER: 40,949  
; REFERENCE/DOCKET NUMBER: 263-PP1R1577US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEFAX: (202) 721-8250  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acid residues  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-318-856A-20

Query Match 34.8%; Score 23; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEEIDY 13  
Db 3 KDELDY 8

Search completed: January 17, 2001, 13:34:39  
Job time: 478 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:34:45 ; Search time 25.93 seconds  
(without alignments)  
21.099 Million cell updates/sec

Title: US-08-765-837-2

Perfect score: 84

Sequence: 1 DEDNQSPRSFQKKTR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 112136

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

A\_Geneseq\_36:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	100.0	16	17 R90946	Factor VIII antige
2	31	36.9	11	9 P81967	Peptide immunogen
3	30	35.7	9	17 R94901	Tastin antigen. H
4	30	35.7	13	19 W39005	Murine DP-3 protei
5	30	35.7	16	18 W29907	Porcine AMPK-alpha
6	29	34.5	15	16 R84179	MKK1 C-terminal pe
7	28	33.3	13	19 W37161	Human TCAK1 antibo
8	28	33.3	14	18 W14614	Human antithrombin
9	27	32.1	9	17 W49604	Human leucocyte an
10	27	32.1	13	19 W87479	HIV-1 mutant env g
11	27	32.1	15	8 P70148	N-terminal of S3-g
12	26	31.0	13	15 R60988	CD4 cytoplasmic do

13	26	31.0	13	19	Y21230	Human bcl2 proto-o
14	26	31.0	15	20	Y13338	Naturally occurin
15	25	29.8	9	16	R70619	HIV(B35)Akv2-39, h
16	25	29.8	12	16	R73639	STAT3 inhibition c
17	25	29.8	13	8	P70536	Antigenic oligopep
18	25	29.8	13	15	R46995	Apolipoprotein B-1
19	25	29.8	13	16	R73630	STAT3 inhibitory p
20	25	29.8	14	16	R69370	N-terminal sequenc
21	25	29.8	15	10	P90817	Non-immunogenic am
22	25	29.8	15	15	R47035	Apolipoprotein B-1
23	25	29.8	15	20	Y50238	Neutrophil-activat
24	24	28.6	7	19	W54304	Human cytohesin-1
25	24	28.6	8	18	W37419	Leptin receptor C-
26	24	28.6	9	18	W17479	Protein kinase C-d
27	24	28.6	9	20	Y52115	GAL4 activation do
28	24	28.6	9	20	Y46140	Immunogenic peptid
29	24	28.6	10	20	Y52103	GAL4 activation do
30	24	28.6	11	19	W36122	Zona pellucida pro
31	24	28.6	11	20	Y29892	Chemokine tumour a
32	24	28.6	11	21	Y68402	Rainbow trout ladd
33	24	28.6	13	15	R49476	EBV tegument prote
34	24	28.6	13	19	W87483	HIV-1 MO/LAT env g
35	24	28.6	13	19	W54680	Peptide from EBV t
36	24	28.6	15	18	W35276	Epitope . . .mpirising
37	24	28.6	15	20	Y32031	Human scieilin pep
38	24	28.6	15	21	Y98859	HLA class II bindi
39	24	28.6	15	21	Y98956	HLA class II bindi
40	24	28.6	15	21	Y68387	Rainbow trout ladd
41	24	28.6	16	15	R66025	Pair 2 pep 3 immun
42	24	28.6	16	15	R66026	Pair 2 pep 4 immun
43	24	28.6	16	16	W11909	Mitochondria loca
44	24	28.6	16	19	W50661	Mitochondria targe
45	23	27.4	5	21	Y61439	Cadherin-7 cell ad

#### ALIGNMENTS

RESULT 1

R90946

ID R90946 standard; peptide: 16 AA.

XX

AC R90946;

XX

XX

DT 09-SEP-1996 (first entry)

XX

DE Factor VIII antigenic peptide corresp. to residues Asp1681-Arg1696.

XX

KW Factor VIII; modification: inhibitor activity; binding; antibody;

KW von Willebrand factor; immune disorder.

XX

OS Synthetic.

XX

PN W09602572-A2.

XX

PD 01-FEB-1996.

XX

PF 14-JUL-1995; 95WO-BE00068.

XX

PR 14-JUL-1994; 94BE-0000666.

XX

XX (CROI-) CROIX ROUGE BELGIQUE.

XX

PI Di Glambattista M, Laub R;

XX

DR WPI; 1996-105861/11.

XX

PT Factor VIII antigenic polypeptide fragments and epitope(s) - also

PT inhibitors of factor VIII and anti-inhibitors, useful for e.g.

PT preventing and treating immune disorders involving inhibition of

XX Factor VIII binding

PS Claim 8; Page 31; 45pp; French.

2 - closed.  
only hit -  
Therein

XX Peptides R90945-64 are derived from the factor VIII protein, esp. from a  
 CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,  
 CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII  
 CC and derived peptides can be used to prevent the activity of inhibitors of  
 CC factor VIII binding to von Willebrand factor, esp. antibodies, thus  
 CC preventing or treating immune disorders.  
 XX  
 SQ Sequence 16 AA;

Query Match 100.0%; Score 84; DB 17; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-08;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEDENQSPRSFQKKTR 16  
 | | | | | | | | | | | | | | | |  
 Db 1 dedengsprsfqkktr 16

RESULT 2  
 P81967  
 ID P81967 standard; protein; 11 AA.  
 AC P81967;  
 XX  
 XX 16-OCT-1990 (first entry)  
 DT  
 XX  
 DE Peptide immunogen for HPV.  
 XX  
 XX Human papilloma virus; immunogen; antibody; antigen; cervical carcinoma;  
 KW HPV-16; early reading frame; E4.  
 XX  
 OS Synthetic.  
 XX  
 XX EP257754-A.  
 PN  
 XX  
 XX 02-MAR-1988.  
 PD  
 XX  
 XX 09-JUL-1987; 87EP-0306061.  
 PF  
 XX  
 XX 10-JUL-1987; 87US-0884184.  
 PR  
 XX  
 XX (STRD ) LELAND STANFORD JUNIOR UNIVERSITY.  
 PA  
 XX  
 XX Schoolnik GK, Palefsky JM;  
 PI  
 XX  
 XX WPI; 1988-057971/09.  
 DR  
 XX  
 XX Synthetic peptide(s) of human papilloma virus - corresp. to amino  
 PT acid sequence region having at least one reverse turn and predicted  
 PT hydrophilicity.  
 XX  
 XX Claim 8; Page 13; 27pp; English.  
 PS  
 XX  
 XX The peptide corresponds to amino acids 48-58 of the E4 early open  
 CC reading frame of HPV-16. It can be used to raise antibodies for  
 CC vaccines or to heighten the immune response to an HPV infection  
 CC already present.  
 CC See also P81955-71.  
 XX  
 XX Sequence 11 AA;

Query Match 36.9%; Score 31; DB 9; Length 11;  
 Best Local Similarity 40.0%; Pred. No. 28;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEDENQSPRS 10  
 | | | | | | | | | |  
 Db 1 ddednsqspet 10

RESULT 3  
 R94901  
 ID R94901 standard; Peptide; 9 AA.  
 XX  
 AC R94901;  
 XX  
 DT 03-JUL-1996 (first entry)  
 XX  
 XX Tactin antigen.  
 DE  
 XX Trophinin; trophinin-assisting protein; tasin; bystin; tasin;  
 KW embryo implantation; infertility; cell adhesion; therapy;  
 KW diagnosis; antibody.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9610414-A1.  
 PN  
 XX  
 PD 11-APR-1996.  
 XX  
 XX 04-OCT-1995; 95WO-US13259.  
 PF  
 XX  
 XX 12-MAY-1995; 95US-0439818.  
 PR  
 XX 04-OCT-1994; 94US-0317522.  
 XX  
 XX (LJOL-) LA JOLLA CANCER RES FOUND.  
 PA  
 XX  
 XX Fukuda MN;  
 PI  
 XX  
 XX WPI; 1996-209192/21.  
 DR  
 XX  
 XX Mammalian trophinin and trophinin-assisting protein - used in  
 PT inhibiting or enhancing embryo implantation, diagnosis of  
 PT infertility and treatment of cancer  
 XX  
 XX Claim 25; Page 80; 106pp; English.  
 PS  
 XX  
 XX A peptide (R94901) corresponding to amino acids 41-49 of human  
 CC trophinin-assisting protein tasin (see also R94900), and  
 CC representing an active fragment of tasin, was used to elise  
 CC antibodies in rabbits. These antibodies were used to detect  
 CC tasin in samples of cells and tissues.  
 XX  
 XX Sequence 9 AA;

Query Match 35.7%; Score 30; DB 17; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
 : | | | | |  
 Db 1 dgenqdp 8

RESULT 4  
 W33005  
 ID W33005 standard; peptide; 13 AA.  
 XX  
 AC W33005;  
 XX  
 XX 22-MAY-1998 (first entry)  
 DT  
 XX  
 DE Murine DP-3 protein derived peptide.  
 XX  
 XX Assaying; cell cycle regulator; E region; DP-3 protein;  
 KW nuclear localisation signal; NLS; mouse; murine.  
 XX  
 XX Mus sp.  
 OS  
 XX WO9743647-A1.  
 PN  
 XX  
 XX 20-NOV-1997.  
 PD

XX 15-MAY-1997; 97WO-GB01324.  
 XX 15-MAY-1996; 96GB-0010195.  
 XX (MEDI-) MEDICAL RES COUNCIL.  
 XX De La Luna S, La Thangue NB;  
 XX WPI; 1998-009053/01.  
 XX Assays for identifying regulators of cell cycle progression -  
 PT comprise expressing a protein having a transcription factor nuclear  
 PT localisation signal and determining the degree of nuclear  
 PT localisation  
 XX Example 1; Page 17; 65pp; English.  
 XX The present sequence was used in the development of a novel assay  
 CC for a putative regulator of cell cycle progression. The assay  
 CC comprises expressing in a cell a protein comprising the E region  
 CC and sufficient C-terminal residues of a DP-3 protein to provide a  
 CC functional nuclear localisation signal (NLS), or the NLS of E2F-1,  
 CC and a marker for nuclear localisation. Then the degree of nuclear  
 CC localisation in the presence and absence of the putative regulator  
 CC is determined. Regulators identified using the assay can be used  
 CC as potential regulators of cell proliferation, or as models for  
 CC rational drug design. Regulation of the NLS may lead to effects  
 CC such as enhanced cell division, blocking of cell cycle progression  
 CC or apoptosis. The regulators may also be used to design other  
 CC candidate regulatory compounds.  
 XX Sequence 13 AA;  
 SQ

Query Match 35.7%; Score 30; DB 19; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 49;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEDENQSPRS 10  
 Db :||| :| |  
 2 eedeedps 11

RESULT 5  
 W29907  
 ID W29907 standard; peptide; 16 AA.  
 XX W29907;  
 AC  
 XX  
 XX 16-MAR-1998 (first entry)  
 DT  
 XX Porcine AMPK-alpha active peptide fragment 2.  
 DE  
 XX 5'-AMP activated protein kinase; AMPK; catalytic subunit;  
 KW protein phosphorylation; cholesterol; fatty acid; pig;  
 KW hormone sensitive lipase; HSL; alpha subunit.  
 XX  
 OS Sus scrofa.  
 XX  
 PN W09725341-A1.  
 XX  
 XX 17-JUL-1997.  
 PD  
 XX 07-JAN-1997; 97WO-US00270.  
 PF  
 XX 08-JAN-1996; 96AU-0007450.  
 PR  
 XX (DART-) DARTMOUTH COLLEGE.  
 PA (SVIN-) ST VINCENTS INST MEDICAL RES.  
 XX  
 XX Kemp BE, Mitchellhill KI, Stapleton DI, Witters LA;  
 PI

DR WPI; 1997-372811/34.  
 XX  
 PT New isolated 5'-AMP-activated protein kinase subunit(s) - used to  
 PT develop products for treating e.g. hypercholesterolaemia, obesity,  
 PT hypoxia, ischaemia, nutrition disorders or diabetes mellitus  
 XX  
 PS Disclosure: Page 36; 63pp; English.  
 XX  
 CC This sequence represent a biologically active peptide derived from the  
 CC 5'-AMP-activated protein kinase (AMPK) catalytic alpha subunit from pig  
 CC liver. This fragment retains at least one of the activities of native  
 CC AMPK-alpha i.e the ability to stimulate phosphorylation of protein  
 CC molecules and the ability to mimic the binding of native AMPK-alpha  
 CC to at least one antibody or ligand molecule. AMPK polypeptides can be  
 CC used to identify compounds which regulate the action of kinases. Such  
 CC fragments can be used to reduce biosynthesis of cholesterol and fatty  
 CC acids. They may also be used to inhibit the release of these molecules  
 CC from intracellular stores by hormone sensitive lipase (HSL). They may  
 CC also be used to reduce cellular malonyl CoA levels and promote the  
 CC beta-oxidation of fatty acids by mitochondria. AMPK-alpha fragments  
 CC could be used in the treatment of e.g. hypercholesterolaemia,  
 CC hyperlipidaemia, obesity, clinical syndromes associated with hypoxia or  
 CC ischaemia (e.g. myocardial infarction) disorders of nutrition and  
 CC diabetes mellitus.  
 XX Sequence 16 AA;  
 SQ

Query Match 35.7%; Score 30; DB 18; Length 16;  
 Best Local Similarity 54.5%; Pred. No. 62;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 DENQSPRSFQK 13  
 Db || :| | ||  
 1 dekesrrlfq 11

RESULT 6  
 R84179  
 ID R84179 standard; Peptide; 15 AA.  
 XX  
 XX R84179;  
 AC  
 XX  
 XX 26-MAR-1996 (first entry)  
 DT  
 XX MKK1 C-terminal peptide.  
 DE  
 XX Megakaryocyte kinase-1; MKK1; cytoplasmic tyrosine kinase;  
 KW cellular signal transduction; leukaemia; thrombocytopenia.  
 KW  
 XX Synthetic.  
 OS  
 XX W09529185-A1.  
 PN  
 XX 02-NOV-1995.  
 PD  
 XX 24-APR-1995; 95WO-US05008.  
 PF  
 XX 21-APR-1995; 95US-0426509.  
 PR  
 XX 22-APR-1994; 94US-0232545.  
 PR  
 XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PA (SUGE-) SUGEN INC.  
 XX  
 XX Gishizky M, Sures I, Ullrich A;  
 PI  
 XX WPI; 1995-382959/49.  
 DR  
 XX New poly:nucleotide(s) encoding megakaryocyte tyrosine kinase(s) -  
 PT used to develop prods. for the treatment and diagnosis of kinase  
 PT related signal transduction abnormalities.  
 XX  
 XX Example 8; Page 41; 82pp; English.  
 PS

xx Synthetic peptides (R84179-80) corresponding to the C-terminal  
 CC regions of novel human megakaryocytic kinases MKK1 (R84181) and MKK2  
 CC (R84182) were used to raise anti-MKK antibodies capable of  
 CC immunoprecipitation of in vitro transcribed and translated MKK1 and  
 CC MKK2 proteins.  
 xx  
 SQ Sequence 15 AA;

Query Match 34.5%; Score 29; DB 16; Length 15;  
 Best Local Similarity 60.0%; Pred. No. 85;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEDENQSPRS 10  
 | | : | | | |  
 Db 3 dadgtsprs 12

RESULT 7  
 W37161  
 ID W37161 standard; peptide; 13 AA.  
 XX  
 AC W37161;  
 XX  
 DT 18-JUN-1998 (first entry)  
 XX  
 DE Human TcAK1 antibody internal C-terminal peptide SEQ ID NO:16.  
 XX  
 KW Human; twenty-five C associated protein kinase 1; TcAK1; detection;  
 KW cancer; cell proliferation; Cdc25; phosphorylation; 14-3-3 protein;  
 KW antibody.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9801756-A1.  
 XX  
 PD 15-JAN-1998.  
 XX  
 PF 03-JUL-1997; 97WO-US11721.  
 XX  
 PR 09-JUL-1996; 96US-0677298.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Piwnicka-Worms H;  
 XX  
 DR WPI; 1998-101198/09.  
 XX  
 PT DNA encoding Twenty-five C Associated protein Kinase 1 - useful for,  
 PT e.g. detecting cancers or disorders of cell proliferation  
 XX  
 PS Claim 12; Page 48; 75pp; English.  
 XX

The present sequence represents a human twenty-five C associated  
 CC protein kinase 1 (TcAK1) antibody internal C-terminal peptide from the  
 CC present invention. The present invention also describes: creating a  
 CC 14-3-3 recognition motif within a TcAK1 substrate by incubating the  
 CC substrate in a cell transformed with the TcAK1 DNA sequence, such that  
 CC a 14-3-3 recognition motif is created in the substrate in a  
 CC phosphorylation-specific manner; and detecting a mutation in a 14-3-3  
 CC protein or a TcAK1 substrate, where the mutation prevents or interferes  
 CC with phosphorylation-specific interaction between the 14-3-3 protein and  
 CC the TcAK1 substrate, by: (a) incubating the 14-3-3 protein and the TcAK1  
 CC substrate which is suspected of being mutated in a cell transformed with  
 CC the TcAK1 DNA sequence such that the 14-3-3 protein and TcAK1 substrate  
 CC interact in a phosphorylation-specific manner; (b) comparing the level  
 CC of interaction with the level obtained with a non-mutant component, and  
 CC (c) diagnosing a mutation in either the 14-3-3 protein or TcAK1 substrate  
 CC if the level of interaction differs significantly from the level obtained  
 CC with the non-mutant components. TcAK1 has a role in mediating interaction  
 CC between 14-3-3 proteins, e.g. human cdc25 phosphorylated on Ser216 which  
 CC creates a 14-3-3 recognition motif and other cellular proteins involved

CC in oncogenesis and signalling events. Measurement of protein binding at  
 CC TcAK1-mediated 14-3-3 recognition sites can be used to detect cancers or  
 CC other disorders of cell proliferation.  
 XX  
 SQ Sequence 13 AA;

Query Match 33.3%; Score 28; DB 19; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEDENQSPRS 10  
 | | : | | | |  
 Db 4 denkeakprs 13

RESULT 8  
 W14614  
 ID W14614 standard; peptide; 14 AA.  
 XX  
 AC W14614;  
 XX  
 DT 09-DEC-1997 (first entry)  
 XX  
 DE Human antithrombin III carboxy-terminal extension.  
 XX  
 KW Human; antithrombin III; mutant; carboxy-terminal; extension;  
 KW sulphated tyrosine; prevention; treatment; thrombosis; disease.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1...3  
 FT /note= "carboxy-terminus of wild type human  
 FT antithrombin III"  
 XX  
 PN JP09071600-A.  
 XX  
 PD 18-MAR-1997.  
 XX  
 PF 06-SEP-1995; 95JP-0228694.  
 XX  
 PR 06-SEP-1995; 95JP-0228694.  
 XX  
 PA (EISA ) EISAI CO LTD.  
 XX  
 DR WPI; 1997-231188/21.  
 XX  
 PT Human antithrombin III mutant with C-terminal peptide containing a  
 PT sulphated tyrosine - useful for treatment and prevention of thrombotic  
 PT disease  
 XX  
 PS Disclosure; Page 6; 11pp; Japanese.  
 XX

CC A human antithrombin III (AT III) mutant, comprising a  
 CC carboxy-terminal sulphated tyrosine peptide extension, i.e. the  
 CC present peptide, can be used to prevent and/or treat thrombotic  
 CC diseases.  
 XX  
 SQ Sequence 14 AA;

Query Match 33.3%; Score 28; DB 18; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEDEN 5  
 | | | | |  
 Db 10 deden 14

RESULT 9  
 W49604



ID W49604 standard; peptide; 9 AA.  
 XX W49604;  
 AC  
 DT 05-JUN-1998 (first entry)  
 XX  
 DE Human leucocyte antigen DQ4 binding peptide #495.  
 XX  
 KW Human leucocyte antigen; HLA-DQ4; combinatorial library;  
 KW autoimmune disease; chronic articular rheumatism.  
 XX  
 OS Synthetic.  
 XX  
 PN JP08151396-A.  
 XX  
 PD 11-JUN-1996.  
 XX  
 PF 28-NOV-1994; 94JP-0292657.  
 XX  
 PR 28-NOV-1994; 94JP-0292657.  
 XX  
 PA (TEIJ ) TEIJIN LTD.  
 XX  
 DR WPI; 1996-329479/33.  
 XX  
 XX HLA-binding oligopeptide and an immuno:regulator contg it - used in  
 PT the treatment of auto:immune disease  
 XX  
 PS Claim 4; Page 53; 61pp; Japanese.  
 XX  
 CC This peptide is an example of a peptide which binds to a human leucocyte  
 CC antigen HLA-DQ4 molecule. The peptide was isolated from a phagemid  
 CC combinatorial library comprising the sequence V05953, by screening with  
 CC an HLA-DQ4 molecule. The peptide is used for the treatment of autoimmune  
 CC disease, or especially for treatment of viral diseases.  
 XX  
 SQ Sequence 9 AA;

Query Match 32.1%; Score 27; DB 17; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 EDENQSPR 9  
 DB 2 edknqrr 9

RESULT 10  
 ID W87479  
 XX W87479 standard; peptide; 13 AA.  
 AC  
 AC W87479;  
 XX  
 DT 11-FEB-1999 (first entry)  
 XX  
 DE HIV-1 mutant env gp41 amino acid fragment in L-2 cell.  
 XX  
 KW Human immunodeficiency virus-1; HIV-1; protease; provirus; L-2 cell;  
 KW mutation; Nef protein; immunogen; apoptosis; treatment; hybridisation;  
 KW cell lysis; chemotherapeutic; toxin; HIV infection.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN W09844945-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 03-APR-1998; 98MO-US06690.  
 XX  
 PR 04-APR-1997; 97US-0043047.  
 XX  
 PA (IMMU-) IMMUNE RESPONSE CORP.

XX Luftig RB;  
 PI  
 DR WPI; 1998-568304/48.  
 DR N-PSDB; V71862.  
 XX  
 PT New nucleic acid encoding mutant or truncated forms of human immune  
 PT deficiency virus proteins - used to generate non-infectious  
 PT particles useful as therapeutic or prophylactic immunogens, also for  
 PT diagnosis  
 XX  
 PS Example 1; Fig 1; 66pp; English.  
 XX  
 CC Sequences W87475 to W87463 represent mutant amino acid sequences of  
 CC the defective human immunodeficiency virus-1 (HIV-1) provirus in L-2  
 CC cells. They represent the portions of pol protease (prot.), vpr, env  
 CC (gp120 and gp41), and nef gene regions that were mutated as compared to  
 CC wild-type HIV-1 in LAI or MO/LAI cells. The invention relates to mutated  
 CC DNA, proteins or fragments of defective HIV-1 provirus in L-2 cell line.  
 CC The mutations result in non-infectious HIV-1 particles. Inactive,  
 CC protease-deficient HIV-1 particles containing at least one of Nef  
 CC truncated protein fragments are used as immunogens, particularly for  
 CC reducing or preventing apoptosis in HIV-1 sero-negative or -positive  
 CC subjects, specifically those with HIV-1 infection, both for prevention  
 CC and treatment. Fragments of nef gene and fragments encoding specific  
 CC mutant Nef proteins are also useful in hybridisation tests for diagnostic  
 CC detection of mutated genes in (lysed) cells or body fluids, while the  
 CC corresponding mutant proteins are detected in immunoassays using  
 CC antibodies. The protease-deficient HIV-1 particles and antibodies,  
 CC optionally attached to a radioisotope, chemotherapeutic agent or toxin,  
 CC can be used to reduce the severity of HIV infections.  
 XX  
 SQ Sequence 13 AA;

Query Match 32.1%; Score 27; DB 19; Length 13;  
 Best Local Similarity 41.7%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DEENQSPRSFQ 12  
 DB 2 eesqqrneq 13

RESULT 11  
 ID P70148  
 XX P70148 standard; peptide; 15 AA.  
 AC  
 AC P70148;  
 XX  
 DT 26-APR-1991 (first entry)  
 XX  
 DE N-terminal of S3-genotype-associated protein.  
 XX  
 KW Self incompatibility; S-protein; pollen tube; style; stigma;  
 KW pollination; gametocide;  
 XX  
 OS Lycopersicon peruvianum.  
 XX  
 PN EP222526-A.  
 XX  
 PD 20-MAY-1987.  
 XX  
 PF 20-OCT-1986; 86EP-0308117.  
 XX  
 PR 21-APR-1986; 86US-0854139.  
 XX  
 PR 29-OCT-1985; 85US-0792435.  
 XX  
 PA (LUBR ) LUBRIZOL GENETICS.  
 XX  
 PI Mau SL, Anderson M, Cornish E, Tregear GW, Crawford RJ;  
 PI Niall HD;  
 XX

DR WPI; 1987-137534/20.  
 XX New DNA sequences of s-genes encoding s-proteins - useful for  
 PT controlling self-incompatibility reaction in self-incompatible  
 PT plants and s-proteins are gametocides etc.  
 XX Disclosure; Fig 4; 42pp; English.  
 XX  
 CC The sequence was used as a basis for the design of probes to iso-  
 CC late DNA encoding the S proteins. The sequence was determined by  
 CC conventional microsequencing techniques performed on protein pur-  
 CC ified from styles dissected from the flowers of the plant.  
 CC The S-proteins control the self-incompatibility reaction. They  
 CC are useful in control of pollen tube growth egas natural game-  
 CC toclides to control, induce or promote self-incompatibility and  
 CC interspecific incompatibility. See also P70149-P70154, N70219  
 CC and N70220.  
 XX  
 XX Sequence 15 AA;  
 SQ

Query Match 32.1%; Score 27; DB 8; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 QSPRSF 11  
 I I I I I  
 Db 10 qxprsf 15

RESULT 12  
 R60988  
 ID R60988 standard; peptide; 13 AA.  
 AC R60988;  
 DT 14-APR-1995 (first entry)  
 XX CD4 cytoplasmic domain fragment.  
 DE  
 XX CD4; T cell; surface antigen; receptor; MHC class II antigen;  
 KW protein-tyrosine kinase; p56lck; TCR/CD3 complex; PI 3-kinase;  
 KW PI 4-kinase; lipid kinase; T cell receptor complex.  
 XX  
 OS Synthetic.  
 XX  
 XX WO9418832-A.  
 PN  
 XX 01-SEP-1994.  
 PD  
 XX 25-FEB-1994; 94WO-0501840.  
 PF  
 XX 26-FEB-1993; 93US-0023915.  
 PR  
 XX (DAND ) DANA FARBER CANCER INST INC.  
 PA  
 XX Cantley L, Kanteti P, Rudd CE;  
 PI  
 XX WPI; 1994-293868/36.  
 DR  
 XX  
 XX Method for inhibiting or reducing signal transduction - utilises  
 PT peptide or corresp. nucleic acid which decreases association of  
 PT PI 3- or 4- kinase with CD4/p56lck  
 PT  
 XX Disclosure; Page 28; 46pp; English.  
 PS  
 XX In order for certain T cells to make an optimal response to antigen,  
 CC it is necessary for the T cell surface antigen CD4 to couple to the  
 CC protein-tyrosine kinase p56lck. (CD4-p56lck is known to associate  
 CC with and functionally synergise with the TCR/CD3 complex.)CD4-p56lck  
 CC complex in T cells associates with two lipid kinases: PI 3-kinase  
 CC and PI 4-kinase, which suggests that these lipid kinases are also  
 CC involved in intracellular signalling via the T cell receptor complex.

CC The interaction of a lipid kinase, such as PI 3-kinase or PI 4-  
 CC kinase, with CD4-p56lck, may be blocked by administering a peptide.  
 CC This peptide may be a fragment of the cytoplasmic domain of CD4  
 CC (eg R60987-R60991), a fragment of p56lck (eg R60992, R60993),  
 CC a fragment of PI 3-kinase (eg R60994, R60995), or a fragment of  
 CC PI 4-kinase. Other proline-rich peptides that bind to SH3 binding  
 CC sequences can also be used, such as the fragment of  
 CC 3BP1 protein that binds to the SH3 of the Abl kinase (R60997), or  
 CC a sequence found in the SOS protein (R60999).  
 XX  
 XX Sequence 13 AA;  
 SQ

Query Match 31.0%; Score 26; DB 15; Length 13;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 QSPRSFQK 13  
 I I I I I  
 Db 5 qcphrfqk 12

RESULT 13  
 Y21230  
 ID Y21230 standard; Protein; 13 AA.  
 XX  
 AC Y21230;  
 XX  
 DT 22-JUL-1999 (first entry)  
 XX  
 XX Human bcl2 proto-oncogene mutant protein fragment 78.  
 DE  
 XX Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;  
 KW frameshift mutation; age-related disease; neurodegenerative disorder;  
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;  
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;  
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;  
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;  
 KW high mobility group protein-C; neuroendocrine specific protein A.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX WO9845322-A2.  
 PN  
 XX 15-OCT-1998.  
 PD  
 XX 02-APR-1998; 98WO-IB00705.  
 PF  
 XX 10-APR-1997; 97US-0043163.  
 PR  
 XX (UYUT-) RIJKSUNIV UTRECHT.  
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.  
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
 XX  
 XX Burbach JPH, Grosveld FG, Van Leeuwen FW;  
 PI  
 XX WPI; 1998-609901/51.  
 DR  
 XX N-PSDB; X75766.  
 DR  
 XX Diagnosing disease by detecting frameshift mutations in RNA or  
 PT corresponding protein mutations - used to diagnose cancer and  
 PT neurological diseases, particularly Alzheimer's disease, and also  
 PT for treatment and prevention with specific ribozymes or wild-type  
 PT RNA  
 XX  
 XX Disclosure; Figure 15; 258pp; English.  
 PS  
 XX This invention describes a novel method for the diagnosis of a disease  
 CC caused by, or associated with, an RNA molecule that has a frameshift

CC mutation. The method is used to diagnose age-related diseases, especially  
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
 CC and many others listed) or susceptibility to these disorders. The method  
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
 CC at an early stage. It is based on the observation that disease may be  
 CC caused by mutations in RNA rather than DNA. The invention describes the  
 CC use of neuronal system RNA molecules, specifically proteins including  
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic  
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HSPF-1, high mobility group  
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 13 AA;

Query Match 31.0%; Score 26; DB 19; Length 13;  
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 QSPRSFQKK 14  
 II I I I I  
 Db 5 qscrifkkl 13

RESULT 14

ID Y13338 standard; peptide; 15 AA.

AC Y13338;

DT 21-JUN-1999 (first entry)

DE Naturally occurring variant of the glutenin epitope Y13221.

KW Gluten; variant; gliadin; glutenin; gluten-sensitive T-cell;  
 KW vaccine; gluten tolerance; gluten-sensitivity; celiac disease; CD;  
 KW celiac sprue; tropical sprue; childhood food allergy;  
 KW dermatitis herpetiformis; DH.

OS Synthetic.

PN EP05518-A1.

XX 31-MAR-1999.

PF 23-SEP-1997; 97EP-0202909.

PR 23-SEP-1997; 97EP-0202909.

XX (ZIEK-) ACAD ZIEKENHUIS LEIDEN.

PA (UYLE-) RIJKSUNIV LEIDEN.

PI Drijfhout JW, Koning F, Kooy-Winkelaar EMC, Van De Wal Y;  
 XX WPI; 1999-192792/17.

DR New antigenic peptides of gluten and methods for isolating them.

PT useful as diagnostic agents and for treatment of gluten sensitivity  
 PT i.e. Celiac disease (CD)

XX Disclosure; Page 46; 58pp; English.

XX Peptides Y13220-13343 represents gluten derived peptides, and their  
 CC variants. The specification describes a method to find and characterize  
 CC peptides that are recognized by an intestinally derived gluten-sensitive  
 CC T-cells. The method comprises establishing and contacting at least one  
 CC gluten-sensitive T-cell clone with a mixture of gluten-derived peptides,  
 CC and fractionating the mixture to select peptides that stimulate the

CC clonal cells from bioactive fractions. Peptides Y13220 (gliadin  
 CC derived peptide) and Y13221 (glutenin derived peptide) were identified  
 CC using these methods, and can be used in pharmaceuticals/medicines  
 CC (vaccines) for inducing tolerance to gluten, or to treat  
 CC gluten-sensitivity, i.e. celiac disease (CD) or celiac sprue, tropical  
 CC sprue, childhood food allergies and dermatitis herpetiformis (DH). The  
 CC peptides are also useful for elimination of a group of gluten-sensitive  
 CC T-cells, and for generating antibodies, T-cell receptors,  
 CC anti-idiotypic B- or T-cells by immunization of a mammal with the  
 CC peptide.

XX Sequence 15 AA;

Query Match 31.0%; Score 26; DB 20; Length 15;  
 Best Local Similarity 40.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 EMQSPRSFQK 13  
 : I I I I I  
 Db 3 qggqprqwqq 12

RESULT 15

R70619

ID R70619 standard; Peptide; 9 AA.

XX R70619;

XX 14-FEB-1996 (first entry)

DE HIV(B35)ARV2-39, human immunodeficiency virus epitope.

XX HLA; human lymphocyte antigen; HIV; human immunodeficiency virus;  
 KW binding peptide; induce killer cell; prevention; treatment; AIDS;  
 KW autoimmune disease syndrome; vaccine.

OS Human immunodeficiency virus.

XX WO9511255-A1.

PD 27-APR-1995.

XX 19-OCT-1994; 94WO-JP01756.

PR 19-OCT-1993; 93JP-0261302.

XX (AJIN) AJINOMOTO CO INC.

PA (AJIN) AJINOMOTO KK.

XX Miwa K, Takiguchi M;

XX WPI; 1995-170188/22.

DR HLA-binding peptide fragments from HIV proteins - induce killing  
 PT cells which target HIV-infected cells and can be incorporated into  
 PT anti-HIV vaccines

XX Example 1; Page 11; 61pp; Japanese.

XX R70619 is a peptide fragment derived from an HIV (Human Immunodeficiency  
 CC Virus) protein and is capable of binding to a human lymphocyte antigen.  
 CC The peptide can induce killer cells which target HIV-infected cells.  
 CC It is also useful in the prevention and treatment of HIV and AIDS.  
 CC Anti-HIV vaccines may incorporate the peptides, or may incorporate a  
 CC vector (such as vaccinia or BCG) containing the peptides.

XX Sequence 9 AA;

Query Match 29.8%; Score 25; DB 16; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 SPRSFQKK 14  
   11111 :  
Db 1 splsfqtr 8

Search completed: January 17, 2001, 13:36:41  
Job time: 116 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:35:50 ; Search time 20.45 Seconds  
(without alignments)  
53.125 Million cell updates/sec

Title: US-08-765-837-2

Perfect score: 84

Sequence: 1 DEDENQSPRSFQKKTR 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 2778

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_66: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	32.1	15	2 PQ0193	stylar glycoprotei
2	25	29.8	15	1 SFRT	scotophobin - rat
3	24	28.6	14	2 A39703	tubulin beta-3 cha
4	23	27.4	10	2 S65387	cytochrome-c oxida
5	23	27.4	13	2 PH0756	T-cell receptor be
6	23	27.4	15	2 A26212	carboxypeptidase B
7	21	25.0	10	2 S70721	heat shock protein
8	21	25.0	13	2 A53608	neurotensin - gulin
9	21	25.0	14	2 S68095	calcium-binding pr
10	21	25.0	15	2 JP0101	fibrinogen alpha c
11	21	25.0	15	2 PA0058	protein QF200022 -
12	21	25.0	15	2 S71300	ICL3 protein - Par
13	21	25.0	16	2 JH0517	insulin-like growt
14	20	23.8	10	2 S71948	matrix metalloprot
15	20	23.8	11	2 S04875	nifS protein - Bra
16	20	23.8	13	2 F61458	Ig kappa chain V-I
17	20	23.8	13	2 S47376	T-cell antigen rec
18	20	23.8	14	2 B34135	DNA-binding protei
19	20	23.8	15	2 I46512	troponin - rabbit
20	20	23.8	15	2 A17340	ribonucleoside-dip
21	20	23.8	15	2 S36889	ribosomal protein
22	20	23.8	15	2 A43839	beta-tubulin germ-
23	20	23.8	15	2 C36198	T-cell receptor be
24	20	23.8	15	2 PH0772	T-cell receptor be
25	20	23.8	16	2 F49039	T-cell receptor be
26	19	22.6	5	2 PT0281	Ig heavy chain CRD
27	19	22.6	8	2 A39892	P element, P cytot
28	19	22.6	8	2 A14683	aspartate transami
29	19	22.6	8	2 S29272	tocopherol-binding

30 19 22.6 11 2 S23373 T-cell receptor al  
31 19 22.6 12 2 S48209 zinc-metallophosph  
32 19 22.6 15 2 B48047 phospholipase C-be  
33 19 22.6 16 2 S05703 homeotic protein c  
34 19 22.6 16 2 S22040 cob protein - comm  
35 19 22.6 16 2 PH0749 T-cell receptor be  
36 18 21.4 7 2 B39040 calsequestrin, fas  
37 18 21.4 10 2 S39392 calpain (LC 3.4.22  
38 18 21.4 10 2 S07202 phyllomedulin - tw  
39 18 21.4 10 2 I40032 trpE protein - Bac  
40 18 21.4 11 2 I54193 Rhesus blood group  
41 18 21.4 15 2 S62620 protein disulfide-  
42 18 21.4 15 2 PQ0192 stylar glycoprotei  
43 18 21.4 15 2 PA0086 protein QF200044 -  
44 18 21.4 15 2 PH1443 T-cell receptor al  
45 18 21.4 15 2 D54226 light-harvesting p

#### ALIGNMENTS

RESULT 1

PQ0193

stylar glycoprotein 3 - Peruvian tomato (fragment)

N: Alternate names: S3-glycoprotein

C: Species: Lycopersicon peruvianum (Peruvian tomato)

C: Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 23-Aug-1997

C: Accession: PQ0193; A28819

R: Mau, S.L.; Williams, E.G.; Atkinson, A.; Anderson, M.A.; Cornish, E.C.; Grayo, B.;  
Planta 169, 184-191, 1986

A: Title: Style proteins of a wild tomato (Lycopersicon peruvianum) associated with ex

A: Reference number: PQ0192

A: Accession: PQ0193

A: Molecule type: protein

A: Residues: 1-15 <MAU>

A: Experimental source: style, strain Mill

R: Anderson, M.A.; Cornish, E.C.; Mau, S.L.; Williams, E.G.; Hoggart, R.; Atkinson, A.;  
J. P.; Crawford, R.J.; Clarke, A.E.

Nature 321, 38-44, 1986

A: Title: Cloning of cDNA for a stylar glycoprotein associated with expression of sel

A: Reference number: A28819

A: Accession: A28819

A: Status: preliminary

A: Molecule type: protein

A: Residues: 1-15 <AND>

C: Keywords: glycoprotein

Query Match 32.1%; Score 27; DB 2; Length 15;  
Best Local Similarity 83.3%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 QSPRSF 11

Db 10 QXPRSF 15

RESULT 2

SFRT

scotophobin - rat

C: Species: Rattus norvegicus (Norway rat)

C: Date: 23-Oct-1981 #sequence\_revision 23-Oct-1981 #text\_change 07-May-1999

C: Accession: A93164; A92757; A01423

R: Ungar, G.; Desiderio, D.M.; Parr, W.

Nature 238, 198-202, 1972

A: Title: Isolation, identification and synthesis of a specific-behaviour-inducing bra

A: Reference number: A93164; MUID:72240741

A: Accession: A93164

A: Molecule type: protein

A: Residues: 1-15 <UNG>

A: Note: Chemical synthesis

R: Desiderio, D.M.; Ungar, G.; White, P.A.

J. Chem. Soc. D Chem. Commun. 1971, 432-433, 1971

A:Title: The use of mass spectrometry in the structural elucidation of scotophobin -- a  
A:Reference number: A92757  
A:Accession: A92757  
A:Molecule type: protein  
A:Residues: 1-15 <DES>  
A:Note: The sequence was determined by mass spectrometry  
R:Stewart, W.W.  
Nature 238, 202-209, 1972  
A:Title: Comments on the chemistry of scotophobin.  
A:Reference number: A93165; MUID:72240742  
A:Contents: annotation; referee's comments on first reference above  
R:Ungar, G.; Desiderio, D.M.; Parr, W.  
Nature 238, 209-210, 1972  
A:Reference number: A93166  
A:Contents: annotation  
A:Note: reply to referee's comments  
R:Wilson, D.  
Nature 320, 313-314, 1986  
A:Title: Scotophobin resurrected as a neuropeptide.  
A:Reference number: A43076; MUID:86175024  
A:Contents: annotation  
A:Note: comparison with other neuropeptides; author presents scotophobin sequence in the  
ating the amino end  
C:Superfamily: scotophobin  
C:Keywords: amidated carboxyl end  
F:15/Modified site: amidated carboxyl end (Tyr) #status experimental

Query Match 29.8%; Score 25; DB 1; Length 15;  
Best Local Similarity 45.5%; Pred. No. 4.4e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 DENQSPRSFQK 13  
| | | | |  
Db 2 DNNQGRKSAQK 12

RESULT 3  
A39703  
tubulin beta-3 chain - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 08-Nov-1991 #sequence\_revision 08-Nov-1991 #text\_change 18-Jun-1993  
C:Accession: A39703  
R:Alexander, J.E.; Hunt, D.F.; Lee, M.K.; Shabanowitz, J.; Michel, H.; Berlin, S.C.; Mac  
Proc. Natl. Acad. Sci. U.S.A. 88, 4685-4689, 1991  
A:Title: Characterization of posttranslational modifications in neuron-specific class II  
A:Reference number: A39703; MUID:91271258  
A:Accession: A39703  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14 <ALE>

Query Match 28.6%; Score 24; DB 2; Length 14;  
Best Local Similarity 44.4%; Pred. No. 6e+02;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DEDENQSPR 9  
: | | | | |  
Db 6 EESAQQGPK 14

RESULT 4  
S65387  
cytochrome-c oxidase (EC 1.9.3.1) chain VII b, cardiac - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 16-Jul-1999  
C:Accession: S65387; S65386  
R:Schaeffer, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.  
Eur. J. Biochem. 230, 235-241, 1995  
A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-term  
A:Reference number: S65372; MUID:95324529

A:Accession: S65387  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <SCH>  
A:Accession: S65386  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <SC2>  
C:Keywords: cardiac muscle; heart; oxidoreductase

Query Match 27.4%; Score 23; DB 2; Length 10;  
Best Local Similarity 83.3%; Pred. No. 6.2e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 SFQKKT 15  
| | | | |  
Db 1 SHQKKT 6

RESULT 5  
PH0756  
T-cell receptor beta chain (I7) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999  
C:Accession: PH0756  
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
J. Exp. Med. 174, 1371-1383, 1991  
A:Title: T cell receptor genes in a series of class I major histocompatibility comple  
allelic exclusion and antigen-specific repertoire.  
A:Reference number: PH0746; MUID:92078846  
A:Accession: PH0756  
A:Molecule type: mRNA  
A:Residues: 1-13 <CAS>  
A:Cross-references: EMBL:X60850; NID:g51482; PIDN:CAA43241.1; PID:g51483  
A:Experimental source: T lymphocyte  
C:Keywords: T-cell receptor

Query Match 27.4%; Score 23; DB 2; Length 13;  
Best Local Similarity 66.7%; Pred. No. 8.2e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 DENQSP 8  
| | | | |  
Db 5 DGNQAP 10

RESULT 6  
A26212  
carboxypeptidase B (EC 3.4.17.2) precursor - marbled lungfish (fragment)  
C:Species: Protopterus aethiopicus (marbled lungfish)  
C:Date: 10-Sep-1987 #sequence\_revision 10-Sep-1987 #text\_change 28-Apr-1993  
C:Accession: A26212  
R:Reeck, G.R.; Neurath, H.  
Biochemistry 11, 3947-3955, 1972  
A:Title: Isolation and characterization of pancreatic procarboxypeptidase B and carbo  
A:Reference number: A90363; MUID:73025047  
A:Accession: A26212  
A:Molecule type: protein  
A:Residues: 1-15 <REE>  
C:Keywords: hydrolase; metallo-carboxypeptidase

Query Match 27.4%; Score 23; DB 2; Length 15;  
Best Local Similarity 80.0%; Pred. No. 9.5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SPRSF 11  
: | | | | |  
Db 3 TPRSFP 7

```
RESULT 7
S70721
heat shock protein C62.5 homolog - Salmonella typhimurium (fragment)
N:Alternate names: high temperature protein G
C:Species: Salmonella typhimurium
C:Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C:Accession: S70721
R:Qi, S.Y.; Li, Y.; Szyroki, A.; Gilles, I.G.; Moir, A.; O'Connor, C.D.
Mol. Microbiol. 17, 523-531, 1995
A:Reference number: S70719; MUID:96100451
A:Accession: S70721
A:Molecule type: protein
A:Residues: 1-10 <QIS>
A:Experimental source: strain SL1344
C:Keywords: ATP binding; heat shock; molecular chaperone

Query Match 25.0%; Score 21; DB 2; Length 10;
Best Local Similarity 44.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 ENQSPRSFQ 12
: | | | |
DB 2 KGOETRGFQ 10

RESULT 8
A53608
neurotensin - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 02-Sep-2000
C:Accession: A53608
R:Shaw, C.; Thim, L.; Conlon, J.M.
FEBS Lett. 202, 187-191, 1986
A:Title: [Ser(7)]neurotensin: isolation from guinea pig intestine.
A:Reference number: A53608; MUID:86248085
A:Accession: A53608
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <SHA>
C:Superfamily: neurotensin
C:Keywords: neuropeptide; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 25.0%; Score 21; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 ENQSPRSF 11
||| | |
DB 4 ENKSRRPY 11

RESULT 9
S68095
calcium-binding protein, 23K - Orchestia cavimana (fragment)
C:Species: Orchestia cavimana
C:Date: 21-Apr-1997 #sequence_revision 21-Apr-1997 #text_change 29-Aug-1997
C:Accession: S68095
R:Luquet, G.; Testeniere, O.; Graf, F.
Biochim. Biophys. Acta 1293, 272-276, 1996
A:Title: Characterization and N-terminal sequencing of a calcium binding protein from th
A:Reference number: S68095; MUID:96202045
A:Accession: S68095
A:Molecule type: protein
A:Residues: 1-14 <LUQ>
C:Keywords: calcium binding

Query Match 25.0%; Score 21; DB 2; Length 14;
Best Local Similarity 40.0%; Pred. No. 1.9e+03;
```

```
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 DEDENQSPRS 10
||| | |
DB 4 DEEDDERRS 13

RESULT 10
JP0101
fibrinogen alpha chain - duck (fragment)
N:Contains: fibrinopeptide A
C:Species: Anas platyrhynchos (domestic duck)
C:Date: 30-Jun-1987 #sequence_revision 28-Dec-1987 #text_change 26-Jan-1996
C:Accession: JP0101
R:Min, Y.; Ping, Z.; Yaoshi, Z.
Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 28, 31-35, 1985
A:Title: Purification and primary structures of duck fibrinopeptides A and B.
A:Reference number: A94238
A:Accession: JP0101
A:Molecule type: protein
A:Residues: 1-15 <MIN>
C:Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology
C:Keywords: blood coagulation; plasma; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 25.0%; Score 21; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 SFQKK 14
|||||
DB 6 SFQKE 10

RESULT 11
PA0058
protein OF200022 - fungus (Fusarium sporotrichioides) (fragment)
C:Species: Fusarium sporotrichioides
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 06-Jun-1997
C:Accession: PA0058
R:Chow, L.P.; Fukaya, N.; Suglura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A:Reference number: PA0051
A:Accession: PA0058
A:Molecule type: protein
A:Residues: 1-15 <CHO>

Query Match 25.0%; Score 21; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSP 8
| | | |
DB 6 ETEXQAP 12

RESULT 12
ICL3 protein - Paramesium tetraurelia (fragment)
C:Species: Paramesium tetraurelia
C:Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-Dec-1999
C:Accession: S71300
R:Madeddu, L.; Klotz, C.; le Caer, J.P.; Beisson, J.
Eur. J. Biochem. 238, 121-128, 1996
A:Title: Characterization of centrin genes in Paramesium.
A:Reference number: S71298; MUID:96248429
A:Accession: S71300
A:Molecule type: protein
A:Residues: 1-15 <MAD>
A:Experimental source: strain d4-2
C:Genetics:
```

A;Genetic code: SGC5

Query Match 25.0%; Score 21; DB 2; Length 15;  
Best Local Similarity 57.1%; Pred. No. 2e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 SPRSFQK 13  
:|:|  
Db 1 APOSAQK 7

#### RESULT 13

JH0517  
insulin-like growth factor-binding protein 4 - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 23-Mar-1995  
C:Accession: JH0517  
R:Coleman, M.E.; Pan, Y.C.E.; Etherton, T.D.  
Biochem. Biophys. Res. Commun. 181, 1131-1136, 1991  
A:Title: Identification and NH2-terminal amino acid sequence of three insulin-like growth  
A:Reference number: JH0515; MUID:92109718  
A:Accession: JH0517  
A:Molecule type: protein  
A:Residues: 1-16 <COL>  
A:Experimental source: serum

Query Match 25.0%; Score 21; DB 2; Length 16;  
Best Local Similarity 37.5%; Pred. No. 2.2e+03;  
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 DEDENQSPRSFOKKTR 16  
||:|:|:|  
Db 1 DEATHXPPPEKLR 16

#### RESULT 14

S71948  
matrix metalloproteinase 3 precursor - bovine (fragments)  
N:Alternate names: MMP-3 protein; Stromelysin precursor MMP-3  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 07-May-1999  
C:Accession: S71948  
R:Arner, E.C.; Pratta, M.A.; Freimark, B.; Lischwe, M.; Trzaskos, J.M.; Magolda, R.L.; W  
Biochem. J. 318, 417-424, 1996  
A:Title: Isothiazolones interfere with normal matrix metalloproteinase activation and in  
A:Reference number: S71948; MUID:96404887  
A:Accession: S71948  
A:Molecule type: protein  
A:Residues: 1-5;6-10 <ARN>  
C:Function:  
A:Description: degrades a wide range of extracellular matrix components, including carti  
A:Note: may be involved in arthritis formation

Query Match 23.8%; Score 20; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 2e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QSPRSF 11  
:|:|  
Db 3 RKPRTF 8

#### RESULT 15

S04875  
nifs protein - Bradyrhizobium japonicum (fragment)  
C:Species: Bradyrhizobium japonicum  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 08-Oct-1999  
C:Accession: S04875  
R:Ebeling, S.  
submitted to the EMBL Data Library, December 1988

A;Reference number: S04873  
A;Accession: S04875  
A:Molecule type: DNA  
A:Residues: 1-11 <EBE>  
A:Cross-references: EMBL:X13691; NID:g39544; PIDN:CAA31382.1; PID:g580775  
C:Genetics:  
A:Gene: nifs  
A:Start codon: GTG

Query Match 23.8%; Score 20; DB 2; Length 11;  
Best Local Similarity 60.0%; Pred. No. 2.2e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 ENOSP 8  
|:|:|  
Db 3 ENRAP 7

Search completed: January 17, 2001, 13:38:36  
Job time: 166 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:36:45 ; Search time 12.32 Seconds  
(without alignments)  
41.940 Million cell updates/sec

Title: US-08-765-837-2  
Perfect score: 84  
Sequence: 1 DEENQSPRSFOKKTR 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 777

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	29.8	15	1 SCOT_RAT	P01159 rattus norv
2	24	28.6	11	1 LADD_ONCMY	P81018 oncorhynch
3	24	28.6	16	1 RL6_VIBPR	Q56715 vibrato prot
4	23	27.4	10	1 COXM_RAT	P80431 rattus norv
5	23	27.4	12	1 RS19_TOBBP	Q56251 tomato big
6	23	27.4	15	1 CBPE_PROAT	P19628 protuberus
7	23	27.4	16	1 PX_RAT	P14630 rattus norv
8	21	25.0	13	1 NEUT_CAVPO	P32560 cavia porce
9	21	25.0	15	1 FIBA_ANAPL	P12801 anas platyr
10	21	25.0	16	1 IBP4_PIG	P24854 sus scrofa
11	20	23.8	8	1 UH09_RAT	P56375 rattus norv
12	20	23.8	11	1 NIFS_BRAJA	P37030 bradyrhizob
13	20	23.8	16	1 LE06_BIOGL	P80745 biomphalari
14	19	22.6	15	1 ITRB_ALEJU	P24927 albizzia ju
15	18	21.4	8	1 UPAL_HUMAN	P30087 homo sapien
16	18	21.4	9	1 UF02_MOUSE	P38640 mus musculu
17	18	21.4	10	1 TKN_PHYBI	P08610 phyllomedus
18	18	21.4	14	1 RS19_PRUAP	Q44160 prunus arme
19	17	20.2	10	1 TKNQ_RANCA	P22690 rana catesb
20	17	20.2	12	1 RS19_CLYEP	Q46490 clover yell
21	17	20.2	15	1 ATP2_PINPS	P81663 pinus pinas
22	16	19.0	8	1 AKH_MELML	P25423 melolontha
23	16	19.0	8	1 ALL7_CARMA	P81809 carcinus ma
24	16	19.0	10	1 MOSQ_CLYJA	P19962 clypeaster
25	16	19.0	11	1 TKN2_UPERU	P08616 uperoleia r
26	16	19.0	12	1 LMT1_LOCFI	P22395 locusta mig
27	16	19.0	12	1 UCRH_MOUSE	P99028 mus musculu
28	16	19.0	13	1 LMT4_LOCFI	P41490 locusta mig
29	16	19.0	13	1 PRC_SALTY	P43669 salmonella
30	16	19.0	15	1 ALB2_TRASC	P81189 trachenys s
31	16	19.0	15	1 GR78_HORSE	P16392 equus cabal
32	16	19.0	15	1 MCA2_RHOOP	P56870 rhodococcus
33	16	19.0	15	1 PSAQ_CUCSA	P42052 cucumis sat

34	16	19.0	16	1 TRYP_FELCA	P81071 felis silve
35	16	19.0	16	1 YMOR_PSEPU	Q02210 pseudomonas
36	15	17.9	8	1 AL15_CARMA	P81818 carcinus ma
37	15	17.9	8	1 ALL3_CXDPO	P82154 cydia pomon
38	15	17.9	8	1 ALL4_CALVO	P41840 cydia pomon
39	15	17.9	8	1 ALL4_CXDPO	P82155 cydia pomon
40	15	17.9	8	1 B44K_PORGI	P81886 porphyromon
41	15	17.9	8	1 FAR3_HOWAM	P41486 homarus ame
42	15	17.9	9	1 FARP_CALSI	P38495 callinectes
43	15	17.9	10	1 RCA_PINPS	P81084 pinus pinas
44	15	17.9	10	1 XYNB_DICB4	P80717 dictyoglomu
45	15	17.9	11	1 TKNA_ONCMY	P28499 oncorhynch

ALIGNMENTS

RESULT 1  
SCOT\_RAT  
ID SCOT\_RAT STANDARD; PRT; 15 AA.  
AC P01159;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE SCOTOPHOBIN.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinus.  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RX MEDLINE=72240741; PubMed=4558348;  
RA Ungar G., Desiderio D.M., Parr W.;  
RT "Isolation, identification and synthesis of a specific-behavior  
RT Inducing brain peptide."  
RL Nature 238:198-202(1972).  
RN [2]  
RP SEQUENCE.  
RA Desiderio D.M., Ungar G., White P.A.;  
RL J. Chem. Soc. Chem. Commun. 432-433(1971).  
RN [3]  
RP REFEREE'S COMMENTS ON REF.1.  
RX MEDLINE=72240742; PubMed=4558349;  
RA Stewart W.W.;  
RT "Comments on the chemistry of scotophobin."  
RL Nature 238:202-209(1972).  
RN [4]  
RP REPLY TO REFEREE'S COMMENTS.  
RA Ungar G., Desiderio D.M., Parr W.;  
RL Nature 238:209-210(1972).  
CC -1- TISSUE SPECIFICITY: BRAIN.  
DR PIR; A01423; SPRT.  
KW Amidation.  
FT MOD\_RES  
SQ SEQUENCE 15 AA; 1582 MW; 5B87696F6C7FD76C CRC64;

Query Match 29.8%; Score 25; DB 1; Length 15;  
Best Local Similarity 45.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 DENQSPRSFOK 13  
Db 2 DNNQGGKSAQQ 12

RESULT 2  
LADD\_ONCMY  
ID LADD\_ONCMY STANDARD; PRT; 11 AA.  
AC P81018;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE LADDERLECTIN (FRAGMENT).

```
OS Onchorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE.
RC TISSUE=BLOOD;
RX MEDLINE=97293418; PubMed=9149391;
RA Jensen L.E., Thiel S., Petersen T.E., Jensenius J.C.;
RT "A rainbow trout lectin with multimeric structure.";
RL Comp. Biochem. Physiol. 116B:385-390(1997).
CC -!- FUNCTION: LECTIN THAT BINDS SEPHAROSE.
CC -!- COFACTOR: CALCIUM IS ESSENTIAL FOR SEPHAROSE BINDING.
CC -!- SUBUNIT: MULTIMERIC.
KW Lectin; Calcium.
FT NON_TER 11
SQ SEQUENCE 11 AA; 1163 MW; 0B26227FF6D45404 CRC64;

Query Match 28.6%; Score 24; DB 1; Length 11;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EDENQSP 8
Dl 1: |||
Dl 3 ENRNQXP 9

RESULT 3
RL6_VIBPR
ID RL6_VIBPR STANDARD; PRT; 16 AA.
AC Q56715;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L6 (FRAGMENT).
GN RPL6.
OS Vibrio proteolyticus (Aeromonas proteolytica).
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15338;
RX Medline=97149305; PubMed=8996113;
RA Satterquist R.A., Smith G.K., Oakley T.H., Lee Y.H., Fox G.E.;
RT "Sequence, overproduction and purification of vibrio proteolyticus
ribosomal protein L18 for in vitro and in vivo studies.";
RL Gene 183.237-242(1996).
CC -!- LOCATION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC LOCATED AT THE AMINOACYL-TRNA BINDING SITE OF THE
CC PEPTIDYLTRANSFERASE CENTER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L6P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38943; AAB41328.1; -.
DR INTERPRO; IPR002358; -.
DR PROSITE; PS00525; RIBOSOMAL_L6.1; PARTIAL.
KW Ribosomal protein; rRNA-binding.
FT NON_TER 1
SQ SEQUENCE 16 AA; 1935 MW; ABC19078DF581B6C CRC64;

Query Match 28.6%; Score 24; DB 1; Length 16;
Best Local Similarity 41.7%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy DENDSPRSFOKK 14
```

```
Db 5 DENVRTKEARKK 16
||| : ||
Db 5 DENVRTKEARKK 16

RESULT 4
COXM_RAT
ID COXM_RAT STANDARD; PRT; 10 AA.
AC P80431;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIIIB (EC 1.9.3.1) (FRAGMENT).
GN COX7B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE.
RC STRAIN=WISTAR; TISSUE=LIVER;
RX MEDLINE=95324529; PubMed=7601105;
RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;
RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and
amino-terminal sequences suggest identity of the fetal heart and the
adult liver isoform.";
RL Eur. J. Biochem. 230:235-241(1995).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1210 MW; CFC70EB771A33326 CRC64;

Query Match 27.4%; Score 23; DB 1; Length 10;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 SFQKKT 15
Dl 1: ||||
Dl 1 SHQKKT 6

RESULT 5
RS19_TOBBP
ID RS19_TOBBP STANDARD; PRT; 12 AA.
AC Q56251;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S19 (FRAGMENT).
GN RPS19.
OS Tomato big bud phytoplasma.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Achleplasmataceae; Phytoplasma.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94350802; PubMed=8071198;
RA Gundersen D.E., Lee I.M., Rehner S.A., Davis R.E., Kingsbury D.T.;
RT "Phylogeny of mycoplasma-like organisms (phytoplasmas): a basis for
their classification.";
RL J. Bacteriol. 176:5244-5254(1994).
CC -!- FUNCTION: PROTEIN S19 FORMS A COMPLEX WITH S13 THAT BINDS STRONGLY
CC TO THE 16S RIBOSOMAL RNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
Qy 10 SFQKKT 15
Dl 1: ||||
Dl 1 SHQKKT 6
```

CC or send an email to license@isb-sib.ch).

CC EMBL: L27004; AAA83952.1; -  
DR INTERPRO: IPR002222; -  
DR PROSITE: PS00323; RIBOSOMAL\_S19; PARTIAL.  
KW Ribosomal protein; rRNA-binding.  
FT NON\_TER 1  
SQ SEQUENCE 12 AA; 1470 MW; 4CC88BE908333AA3 CRC64;

Query Match 27.4%; Score 23; DB 1; Length 12;  
Best Local Similarity 40.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 NQSPRSFQKK 14  
I I I I I  
DB 3 NKKDKKQKK 12

RESULT 6  
CBPB\_PROAT STANDARD; PRT; 15 AA.  
AC P19628;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE CARBOXYPEPTIDASE B (EC 3.4.17.2) (FRAGMENT).  
OS Protopterus aethiopicus (Marbled lungfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Dipnoi; Lepidosireniformes; Protopteridae; Protopterus.  
[1]  
RN NTS.  
RP SEQUENCE.  
RC TISSUE-PANCREAS;  
RX MEDLINE-73025047; PubMed-5079891;  
RA Reek G.R., Neurath H.;  
RT "Isolation and characterization of pancreatic procarboxypeptidase B  
and carboxypeptidase B of the African lungfish."  
RL Biochemistry 11:3947-3955(1972).  
CC -1- CATALYTIC ACTIVITY: PEPTIDYL-L-LYSINE/ARGININE + H(2)O - PEPTIDE +  
L-LYSINE/ARGININE.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M14; ALSO KNOWN AS THE  
ZINC CARBOXYPEPTIDASE FAMILY.  
CC PIR; A26212; A26212.  
DR MEROPS: M14.003; -  
DR INTERPRO: IPR000834; -  
DR PROSITE: PS00132; CARBOXYPEPT\_2N\_1; PARTIAL.  
DR PROSITE: PS00133; CARBOXYPEPT\_2N\_2; PARTIAL.  
KW Hydrolase; Carboxypeptidase; Zinc; zymogen.  
FT PROPEP 1 >15  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1749 MW; 124C910D937BED65 CRC64;

Query Match 27.4%; Score 23; DB 1; Length 15;  
Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 SPRSF 11  
I I I I I  
DB 3 TPRSF 7

RESULT 7  
PX\_RAT STANDARD; PRT; 16 AA.  
AC P14630;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-JUL-1993 (Rel. 26, Last annotation update)  
DE PROTEIN PX (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

RN SEQUENCE.  
RP MEDLINE-90122905; PubMed-2297521;  
RX Blatter M.-C., James R.W., Borghini I., Martin B.M.,  
RA Hochstrasser A.-C., Pometta D.;  
RT "A novel high-density lipoprotein particle and associated protein in  
rat plasma."  
RL Biochim. Biophys. Acta 1042:19-27(1990).  
CC -1- FUNCTION: UNKNOWN.  
CC -1- TISSUE SPECIFICITY: PLASMA PROTEIN ASSOCIATED WITH HDL.  
KW Plasma; HDL. 16 16  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1740 MW; 79836ECDAE9FBA29 CRC64;

Query Match 27.4%; Score 23; DB 1; Length 16;  
Best Local Similarity 45.5%; Pred. No. 3.8e+02;  
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 NQSPRSFQKK 15  
I I I I I  
DB 4 NQXPEXSQLT 14

RESULT 8  
NEUT\_CAVPO STANDARD; PRT; 13 AA.  
AC P32560;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE NEUROTENSIN (NT).  
GN NTS.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
RN [1]  
RP SEQUENCE.  
RC TISSUE-SMALL INTESTINE;  
RX MEDLINE-86248085; PubMed-3087775;  
RA Shaw C., Thim L., Conlon J.M.;  
RT "[Ser7]neurotensin: isolation from guinea pig intestine."  
RL FEBS Lett. 202:187-192(1986).  
CC -1- FUNCTION: SMOOTH MUSCLE-CONTRACTING PEPTIDE.  
DR PIR; A53608; A53608.  
KW Vasoactive.  
FT MOD\_RES 1 1  
SQ SEQUENCE 13 AA; 1680 MW; 4C8314644C4115B3 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 13;  
Best Local Similarity 50.0%; Pred. No. 6.5e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 ENQSPRSF 11  
I I I I I  
DB 4 ENKSRRPY 11

RESULT 9  
FIBA\_ANAPL STANDARD; PRT; 15 AA.  
AC P12801;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE FIBRINOPEPTIDE A.  
OS Anas platyrhynchos (Domestic duck).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE-85168193; PubMed-3983613;

RA Min Y., Ping Z., Yaoshi Z.;  
 RT "Purification and primary structures of duck fibrinopeptides A and B";  
 RL Sci. Sin., B, Chem. Biol. Agric. Med. Earth Sci. 28:31-35(1985).  
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR PIR: JP0101; JP0101.  
 KW Blood coagulation; Plasma.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1580 MW; D78A51FF8B840373 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 15;  
 Best Local Similarity 80.0%; Pred. No. 7.6e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 SFQKK 14  
 IIII:  
 DB 6 SFQKE 10

RESULT 10  
 IBP4\_PIG STANDARD; PRT; 16 AA.  
 AC P24854;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 (IGFBP-4) (IBP-4)  
 DE (IGF-BINDING PROTEIN 4) (FRAGMENT).  
 GN IGFBP4.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE-92109718; PubMed-1722398;  
 RA Coleman M.E., Pan Y.-C.E., Etherton T.D.;  
 RT "Identification and NH2-terminal amino acid sequence of three  
 RT insulin-like growth factor-binding proteins in porcine serum.";  
 RL Biochem. Biophys. Res. Commun. 181:1131-1136(1991).  
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS  
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
 CC PROMOTING EFFECTS OF THE IGFS ON CELL CULTURE. THEY ALTER THE  
 CC INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
 CC PROTEIN FAMILY.  
 DR PIR: JH0517; JH0517.  
 DR INTERPRO: IPR000716; -.  
 DR INTERPRO: IPR000867; -.  
 DR PROSITE: PS00222; IGF-BINDING; PARTIAL.  
 DR PROSITE: PS00484; THYROGLOBULIN\_1; PARTIAL.  
 KW Growth factor binding.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1799 MW; 40988840096655E2 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 16;  
 Best Local Similarity 37.5%; Pred. No. 8.2e+02;  
 Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 DEDENQSPRSEFKKTR 16  
 ||~: |||:|

DB 1 DEAHXPPPEEKLR 16  
 RESULT 11  
 UH09\_RAT STANDARD; PRT; 8 AA.  
 AC P56575;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF HEART TISSUE (SPOT P9) (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-WISTAR; TISSUE=HEART;  
 RA Li X.-P., Pleissner K.-P., Scheler C., Reglitz-Zagrosek V., Salikov J.,  
 RA Jungblut P.R.;  
 RL Submitted (SEP-1998) to the SWISS-PROT data bank.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 8.9, ITS MW IS: 42 KDA.  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 1029 MW; 9E0775A6C4140B06 CRC64;

Query Match 23.8%; Score 20; DB 1; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSP 8  
 :: |||  
 DB 1 QERRQSP 7

RESULT 12  
 NIFS\_BRAJA STANDARD; PRT; 11 AA.  
 AC P37030;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE NIFS PROTEIN (FRAGMENT).  
 GN NIFS.  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Bradyrhizobium group; Bradyrhizobium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ebeling S.;  
 RL Submitted (DEC-1988) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: NOT KNOWN YET.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: X13691; CAA31982.1; -.  
 DR PIR: S04875; S04875.  
 DR INTERPRO: IPR000192; -.  
 DR PROSITE: PS00595; AA\_TRANSFER\_CLASS\_5; PARTIAL.  
 KW Nitrogen fixation; Pyridoxal phosphate.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1308 MW; 09F0BB69476DC404 CRC64;

Query Match 23.8%; Score 20; DB 1; Length 11;  
Best Local Similarity 60.0%; Pred. No. 7.9e+02;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 EDNOSP 8  
DB 3 ENRAP 7

RESULT 13  
LE06\_BIOGL STANDARD; PRT; 16 AA.  
AC P80745;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HEMOLYMPH 65 KDA LECTIN BG06 (FRAGMENT).  
GN BG06.  
OS Biomphalaria glabrata (Bloodfluke planorb).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
OC Planorbidae; Biomphalaria.  
RN [1]  
RP SEQUENCE.  
RC STRAIN=M-LINE; TISSUE=HEMOLYMPH;  
RX MEDLINE=97385165; PubMed=9238039;  
RA Adema C.M., Hertel L.A., Miller R.D., Loker E.S.;  
RT "A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:8691-8696(1997).  
CC -1- FUNCTION: BINDS AND PRECIPITATES ANTIGENS OF THE PARASITE ECHINOSTOMA PARAENSEI.  
CC -1- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH.  
CC -1- INDUCTION: BY INFECTION.  
KW Lectin.  
FT NON\_TER 1 1  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1964 MW; A1665754589EP82C CRC64;

Query Match 23.8%; Score 20; DB 1; Length 16;  
Best Local Similarity 42.9%; Pred. No. 1.2e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DEDENQS 7  
DB 8 DNDQQN 14

RESULT 14  
ITRB\_ALBJU STANDARD; PRT; 15 AA.  
AC P24927;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE TRYPSIN INHIBITOR B CHAIN (FRAGMENT).  
OS Albizzia julibrissin (Silk tree).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
OC Fabales; Fabaceae; Mimosoideae; Albizia.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=SEED;  
RX MEDLINE=80115605; PubMed=528539;  
RA Odani S., Ono T., Ikenaka T.;  
RT "Proteinase inhibitors from a mimosoideae legume, Albizzia julibrissin. Homologues of soybean trypsin inhibitor (Kunitz).";  
RL J. Biochem. 86:1795-1805(1979).  
CC -1- FUNCTION: INHIBITS TRYPSIN AND ALPHA-CHYMOTRYPSIN.  
CC -1- SUBUNIT: HETERODIMER OF AN "A" AND A "B" CHAIN LINKED BY A DISULFIDE BOND.  
CC -1- SIMILARITY: TO BETA CHAINS OF SOME OTHER LEGUMINOUS KUNITZ-TYPE INHIBITORS.

KW Serine protease inhibitor.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1705 MW; 53165F7E9C45B4D0 CRC64;

Query Match 22.6%; Score 19; DB 1; Length 15;  
Best Local Similarity 75.0%; Pred. No. 1.6e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEDE 4  
DB 12 DDDE 15

RESULT 15  
UPAL\_HUMAN STANDARD; PRT; 8 AA.  
AC P30087;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 2) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=PLASMA;  
RX MEDLINE=93092937; PubMed=1459097;  
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquall C.,  
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
RA Hochstrasser D.F.;  
RT "Plasma protein map: an update by microsequencing.";  
RL Electrophoresis 13:707-714(1992).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 4.9, ITS MW IS: 65 KDA.  
DR SWISS-2DPAGE; P30087; HUMAN.  
FT NON\_TER 1 1  
FT NON\_TER 8 8  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 21.4%; Score 18; DB 1; Length 8;  
Best Local Similarity 37.5%; Pred. No. 8.9e+04;  
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
DB 1 DQESNVPR 8

Search completed: January 17, 2001, 13:39:28  
Job time: 163 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:36:15 ; Search time 32.75 seconds  
(without alignments)  
57.262 Million cell updates/sec

Title: US-08-765-837-2

Perfect score: 84

Sequence: 1 EDENOSPRSFRKTR 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 2612

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_15:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	37	44.0	16	2	Q47603
2	26	31.0	16	12	Q79455
3	26	31.0	16	12	Q79461
4	25	29.8	10	12	Q66190
5	25	29.8	13	11	Q63940
6	24	28.6	8	11	Q9JLD7
7	24	28.6	15	11	Q35411
8	23	27.4	11	11	Q9JLE6
9	22	26.2	13	4	Q9Y674
10	22	26.2	15	13	Q9PS10
11	22	26.2	16	10	Q9S8H7
12	21	25.0	15	8	P82135
13	21	25.0	16	10	P82161
14	21	25.0	16	11	Q9QUW8
15	21	25.0	16	12	Q83960
16	20	23.8	9	12	Q67605
17	20	23.8	9	12	Q67606
18	20	23.8	11	2	Q9RF22
19	20	23.8	15	2	Q9R546

20	20	23.8	15	2	Q9R4U8	Q9R4U8 actin(tobact
21	20	23.8	15	10	Q40562	Q40562 nicotiana t
22	20	23.8	15	10	P82331	P82331 pisum sativ
23	20	23.8	15	11	Q9QV17	Q9QV17 rattus sp.
24	19.5	23.2	13	12	Q81761	Q81761 hepatitis c
25	19.5	23.2	13	12	Q81762	Q81762 hepatitis c
26	19.5	23.2	13	12	Q81763	Q81763 hepatitis c
27	19.5	23.2	13	12	Q81764	Q81764 hepatitis c
28	19.5	23.2	13	12	Q81765	Q81765 hepatitis c
29	19.5	23.2	13	12	Q81767	Q81767 hepatitis c
30	19.5	23.2	13	12	Q81768	Q81768 hepatitis c
31	19.5	23.2	13	12	Q81772	Q81772 hepatitis c
32	19.5	23.2	13	12	Q81773	Q81773 hepatitis c
33	19.5	23.2	13	12	Q81774	Q81774 hepatitis c
34	19.5	23.2	13	12	Q81775	Q81775 hepatitis c
35	19.5	23.2	13	12	Q81777	Q81777 hepatitis c
36	19.5	23.2	13	12	Q81778	Q81778 hepatitis c
37	19.5	23.2	13	12	Q81779	Q81779 hepatitis c
38	19.5	23.2	13	12	Q81780	Q81780 hepatitis c
39	19.5	23.2	13	12	Q81782	Q81782 hepatitis c
40	19.5	23.2	13	12	Q81783	Q81783 hepatitis c
41	19.5	23.2	13	12	Q81784	Q81784 hepatitis c
42	19.5	23.2	13	12	Q81785	Q81785 hepatitis c
43	19.5	23.2	13	12	Q81786	Q81786 hepatitis c
44	19.5	23.2	13	12	Q81787	Q81787 hepatitis c
45	19.5	23.2	13	12	Q81788	Q81788 hepatitis c

#### ALIGNMENTS

RESULT 1  
Q47603 PRELIMINARY; PRT; 16 AA.  
ID Q47603  
AC Q47603;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE C (FRAGMENT).  
GN C.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91139577; PubMed=1995588;  
RA Tao T., Bourne J.C., Blumenthal R.M.;  
RT "A family of regulatory genes associated with type II restriction-  
modification systems".  
RL J. Bacteriol. 173:1367-1375(1991).  
DR EMBL; M63621; AAA24559.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1901 MW; 45E8EA71129D7FCE CRC64;

Query Match 44.0%; Score 37; DB 2; Length 16;  
Best Local Similarity 77.8%; Pred. No. 5.3;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EDENOSPRS 10  
| | | | |  
DB 7 EHENSFPRS 15

RESULT 2  
Q79455 PRELIMINARY; PRT; 16 AA.  
ID Q79455  
AC Q79455;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE LIGHT-CHAIN COMPLEMENTARITY-DETERMINING REGION 3 MRNA (CLONE 14).

DE PARTIAL CDS (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92052225; PubMed=1719545;  
 RA Burton D.R., Barbas C.F. III, Persson M.A., Koenig S., Chanock R.M.,  
 RA Lerner R.A.;  
 RT "A large array of human monoclonal antibodies to type 1 human  
 RT immunodeficiency virus from combinatorial libraries of asymptomatic  
 RT seropositive individuals.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10134-10137(1991).  
 DR EMBL; M80720; AAA44780.1; -;  
 DR HSP; P01607; IREI. 1  
 FT NON\_TER 16  
 FT NON\_TER 16  
 SQ SEQUENCE 16 AA; 1883 MW; 9015605C32B3D65E CRC64;

Query Match 31.0%; Score 26; DB 12; Length 16;  
 Best Local Similarity 57.1%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 NQSPRSF 11  
 I : I I I I  
 Db 7 NSAPRTF 13

RESULT 3  
 Q79461 ID 079461 PRELIMINARY; PRT; 16 AA.  
 AC 079461;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE LIGHT-CHAIN COMPLEMENTARITY-DETERMINING REGION 3 MRNA  
 DE (CLONE 31), PARTIAL CDS (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92052225; PubMed=1719545;  
 RA Burton D.R., Barbas C.F. III, Persson M.A., Koenig S., Chanock R.M.,  
 RA Lerner R.A.;  
 RT "A large array of human monoclonal antibodies to type 1 human  
 RT immunodeficiency virus from combinatorial libraries of asymptomatic  
 RT seropositive individuals.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10134-10137(1991).  
 DR EMBL; M80726; AAA44786.1; -;  
 FT NON\_TER 16  
 FT NON\_TER 16  
 SQ SEQUENCE 16 AA; 1883 MW; 9015605C32B3D65E CRC64;

Query Match 31.0%; Score 26; DB 12; Length 16;  
 Best Local Similarity 57.1%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 NQSPRSF 11  
 I : I I I I  
 Db 7 NSAPRTF 13

RESULT 4  
 Q66190 ID 066190 PRELIMINARY; PRT; 10 AA.  
 AC 066190;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL PROTEIN (FRAGMENT).

OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11120;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=41;  
 RX MEDLINE=87021475; PubMed=2429473;  
 RA Niesters H.G., Lenstra J.A., Spaan W.J., Zijderveld A.J.,  
 RA Bieumink-Pluym N.M., Hong F., van Scharrenburg G.J., Horzinek M.C.,  
 RA der Zeijst B.A.;  
 RT "The peplomer protein sequence of the M41 strain of coronavirus IBV  
 RT and its comparison with Beaudette strains.";  
 RL Virus Res. 5:253-263(1986).  
 DR EMBL; M21883; AAA66576.1; -;  
 KW Hypothetical protein.  
 FT NON\_TER 10  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1136 MW; CF2510D5A1B775A6 CRC64;

Query Match 29.8%; Score 25; DB 12; Length 10;  
 Best Local Similarity 83.3%; Pred. No. 4.4e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 QSPRSF 11  
 I I I I I  
 Db 3 QSPTSF 8

RESULT 5  
 Q63940 ID 063940 PRELIMINARY; PRT; 13 AA.  
 AC 063940;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE TROPONIN T (FRAGMENT).  
 DE Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=101116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94140941; PubMed=8308072;  
 RA Morgan M.J., Earnshaw J.C., Dhoot G.K.;  
 RT "Novel developmentally regulated exon identified in the rat fast  
 RT skeletal muscle troponin T gene.";  
 RL J. Cell Sci. 106:903-908(1993).  
 DR EMBL; S68786; CAB32909.1; -;  
 FT NON\_TER 13  
 FT NON\_TER 13  
 SQ SEQUENCE 13 AA; 1601 MW; 7EFB4735373165A4 CRC64;

Query Match 29.8%; Score 25; DB 11; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 5.7e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 SPRSFOKKTR 16  
 I I I I : I I  
 Db 3 SPRSRGRRTR 12

RESULT 6  
 Q9JLD7 ID 09JLD7 PRELIMINARY; PRT; 8 AA.  
 AC 09JLD7;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE P53 TUMOR SUPPRESSOR (FRAGMENT).  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;



RT-PCR analysis of the rat thioredoxin reductase 1 gene.";

○

Query Match 26.2% Score 22; DB 13; Length 15;

Best Local Similarity 80.0%; Pred. No. 2.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 ENQSP 8  
DB 11 ENQRP 15

RESULT 11  
Q9S8H7 PRELIMINARY; PRT; 16 AA.  
AC Q9S8H7; (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-JUN-2000 (TREMREL. 14, Last annotation update)  
DE BETA-GLUCOSIDASE (FRAGMENT).  
OS Pinus contorta (Shore pine) (Lodgepole pine).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Coniferopsida; Coniferales; Pinaceae; Pinus.  
OX NCBI\_TaxID=3339;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=95241620; PubMed=7724669;  
RA Dharmawardhana D.P., Ellis B.E., Carlson J.E.;  
RT "A beta-glucosidase from lodgepole pine xylem specific for the lignin  
precursor coniferin."  
RL Plant Physiol. 107:331-339(1995).  
SQ SEQUENCE 16 AA; 1812 MW; A7072AD36067C086 CRC64;

Query Match 26.2%; Score 22; DB 10; Length 16;

Best Local Similarity 44.4%; Pred. No. 2.4e+03;  
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 DENQSPRSF 11  
DB 3 DRNNFPSDF 11

RESULT 12  
P82135 PRELIMINARY; PRT; 15 AA.  
AC P82135;  
DT 01-JUN-2000 (TREMREL. 14, Created)  
DT 01-JUN-2000 (TREMREL. 14, Last sequence update)  
DT 01-JUN-2000 (TREMREL. 14, Last annotation update)  
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S18 ALPHA (FRAGMENT).  
OS Spinacia oleracea (Spinach).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;  
OC Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV. ALVARO; TISSUE=LEAF;  
RA Yamauchi K., von Knoblauch K., Subramanian A.R.;  
RT "Identification of all the proteins in the small subunit of an  
organelle (chloroplast) ribosome."  
RL J. Biol. Chem. 0:0-0(2000).  
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.  
CC -|- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -|- MISCELLANEOUS: S18 ALPHA AND BETA FORMS DIFFER IN PI. S18 BETA  
FORM IS THE LEAST BASIC FORM.  
CC -|- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.  
DR INTERPRO: IPR001648; -  
KW PROSITE; PS00057; RIBOSOMAL\_S18; PARTIAL.  
DR Ribosomal protein; Chloroplast.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1910 MW; C29396F592403B60 CRC64;

Query Match 25.0%; Score 21; DB 8; Length 15;

Best Local Similarity 41.7%; Pred. No. 3.4e+03;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 NQSPRSFQKKTR 16  
DB 2 DSKRPFIKSRR 13

RESULT 13  
P82161 PRELIMINARY; PRT; 16 AA.  
AC P82161;  
DT 01-JUN-2000 (TREMREL. 14, Created)  
DT 01-JUN-2000 (TREMREL. 14, Last sequence update)  
DT 01-JUN-2000 (TREMREL. 14, Last annotation update)  
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S10 BETA (FRAGMENT).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;  
OC Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV. ALVARO; TISSUE=LEAF;  
RA Yamauchi K., von Knoblauch K., Subramanian A.R.;  
RT "Identification of all the proteins in the small subunit of an  
organelle (chloroplast) ribosome."  
RL J. Biol. Chem. 0:0-0(2000).  
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.  
CC -|- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -|- MISCELLANEOUS: THERE ARE TWO FORMS DUE MAYBE TO A DIFFERENT  
TRANSIT PEPTIDE CLEAVAGE.  
CC -|- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.  
DR INTERPRO: IPR001848; -  
DR PROSITE; PS00361; RIBOSOMAL\_S10; PARTIAL.  
KW Ribosomal protein; Chloroplast.  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1666 MW; 4B7BE4270FF59B27 CRC64;

Query Match 25.0%; Score 21; DB 10; Length 16;

Best Local Similarity 66.7%; Pred. No. 3.6e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 SPSPSF 12  
DB 9 SPDSFE 14

RESULT 14  
Q9QUW8 PRELIMINARY; PRT; 16 AA.  
AC Q9QUW8;  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-JUN-2000 (TREMREL. 14, Last annotation update)  
DE PROTEIN ISOMERASE-RELATED PROTEIN PRECURSOR 71.5 KDA ISOFORM  
(FRAGMENT).  
OS Cavia (guinea pigs).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae.  
OX NCBI\_TaxID=10140;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96070119; PubMed=8535285;  
RA Bonifacio M.D., Steeves T., Saunders D.M., Sinosich M.J.;  
RT "Isolation of ERp72 from guinea pig term placenta using heparin  
Sepharose affinity chromatography."  
RL Biochem Mol Biol Int. 36:1143-1152(1995).  
SQ SEQUENCE 16 AA; 1880 MW; 21B44A5F5767CB17 CRC64;

Query Match 25.0%; Score 21; DB 11; Length 16;

Best Local Similarity 50.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEDENQ 6  
|:|:|  
Db 1 DDDEEE 6

RESULT 15

Q83960 PRELIMINARY; PRT; 16 AA.  
AC Q83960;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE INFLUENZA A/ANN ARBOR/6/60 (H2N2), NON-STRUCTURAL PROTEIN  
DE (Seg 6), COOH TERMINUS OF NS1 (FRAGMENT).  
OS Influenza A virus.  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza virus A and B group.  
OX NCBI\_TaxID=11320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83303830; PubMed=6612993;  
RA Parvin J.D., Young J.F., Palese P.;  
RT "Nonsense mutations affecting the lengths of the NS1 nonstructural  
RT proteins of influenza A virus isolates.";  
RL Virology 128:512-517(1993).  
DR EMBL; K00962; AAA43516.1; -.  
DR INTERPRO; IPR000256; -.  
DR PFAM; PF00600; Flu\_NS1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1811 MW; 23F434B5AEFFAD93 CRC64;

Query Match 25.0%; Score 21; DB 12; Length 16;  
Best Local Similarity 66.7%; Pred. No. 3.6e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 DENQSP 8  
|:|:|  
Db 6 DENGRR 11

Search completed: January 17, 2001, 13:39:12  
Job time: 177 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:35:00 ; Search time 18.64 Seconds  
(without alignments)  
15.414 Million cell updates/sec

Title: US-08-765-837-2

Perfect score: 84

Sequence: 1 DEDENSPRSFQKKTR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 81463

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*

1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2.6/ptodata/2/iaa/6\_COMB.pep.\*

4: /cgn2.6/ptodata/2/iaa/PTUS\_COMB.pep.\*

5: /cgn2.6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	35	41.7	13	5	5258287-29	Patent No. 5258287
2	35	41.7	13	5	5332671-10	Patent No. 5332671
3	31	36.9	11	1	US-07-909-122-13	Sequence 13, Appl
4	31	36.9	11	2	US-08-934-915-189	Sequence 189, Appl
5	30	35.7	9	1	US-08-317-522A-11	Sequence 11, Appl
6	30	35.7	9	1	US-08-439-818A-11	Sequence 11, Appl
7	30	35.7	9	2	US-08-751-965-11	Sequence 11, Appl
8	30	35.7	9	2	US-08-738-975-11	Sequence 11, Appl
9	30	35.7	9	2	US-08-728-626-11	Sequence 11, Appl
10	30	35.7	9	3	US-08-808-599A-11	Sequence 11, Appl
11	30	35.7	13	2	US-08-723-415B-16	Sequence 16, Appl
12	30	35.7	13	3	US-09-189-627A-16	Sequence 16, Appl
13	30	35.7	16	3	US-09-101-146-14	Sequence 14, Appl
14	29	34.5	11	5	5240706-13	Patent No. 5240706
15	28	33.3	13	2	US-08-677-298-16	Sequence 16, Appl
16	28	33.3	16	3	US-08-467-023-246	Sequence 246, Appl
17	28	33.3	16	5	5196511-24	Patent No. 5196511
18	26	31.0	13	4	PCT-US94-01840-2	Sequence 2, Appl
19	26	31.0	13	5	5432076-2	Patent No. 5432076
20	26	31.0	14	1	US-08-709-558A-5	Sequence 5, Appl
21	25	29.8	9	1	US-08-615-181-105	Sequence 105, Appl
22	25	29.8	13	2	US-08-480-190-230	Sequence 230, Appl
23	25	29.8	13	2	US-08-488-379-230	Sequence 230, Appl
24	25	29.8	13	4	PCT-US93-07545-230	Sequence 230, Appl
25	25	29.8	14	1	US-08-362-567-1	Sequence 1, Appl
26	25	29.8	15	1	US-08-218-025A-123	Sequence 123, Appl
27	25	29.8	15	1	US-08-240-514-29	Sequence 29, Appl
28	25	29.8	15	2	US-08-612-302A-29	Sequence 29, Appl

29	25	29.8	15	2	US-08-480-190-270	Sequence 270, Appl
30	25	29.8	15	2	US-08-488-379-270	Sequence 270, Appl
31	25	29.8	15	4	PCT-US93-07545-270	Sequence 270, Appl
32	24	28.6	9	1	US-08-594-447-38	Sequence 38, Appl
33	24	28.6	9	1	US-08-541-964-37	Sequence 37, Appl
34	24	28.6	9	2	US-08-665-647-52	Sequence 52, Appl
35	24	28.6	11	1	US-08-355-888A-17	Sequence 17, Appl
36	24	28.6	11	2	US-08-693-697-17	Sequence 17, Appl
37	24	28.6	11	3	US-08-693-696-17	Sequence 17, Appl
38	24	28.6	13	2	US-08-480-190-186	Sequence 186, Appl
39	24	28.6	13	2	US-08-488-379-186	Sequence 186, Appl
40	24	28.6	13	4	PCT-US93-07545-186	Sequence 186, Appl
41	24	28.6	14	1	US-08-218-025A-28	Sequence 28, Appl
42	24	28.6	15	1	US-07-859-291C-26	Sequence 26, Appl
43	24	28.6	15	1	US-08-388-267C-18	Sequence 18, Appl
44	24	28.6	16	1	US-08-480-463-7	Sequence 7, Appl
45	24	28.6	16	2	US-08-480-445A-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1  
5258287-29  
; Patent No. 5258287  
; APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.  
; TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION  
; OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BPS3  
; NUMBER OF SEQUENCES: 58  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/171,623  
; FILING DATE: 22-MAR-1988  
; SEQ ID NO: 29:  
; LENGTH: 13  
5258287-29

Query Match 41.7%; Score 35; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SFQKKTR 16  
Db 1 SFQKKTR 7

RESULT 2  
5332671-10  
; Patent No. 5332671  
; APPLICANT: FERRARA, NAPOLEONE; LEUNG, DAVID W.H.  
; TITLE OF INVENTION: PRODUCTION OF VASCULAR ENDOTHELIAL CELL  
; GROWTH FACTOR AND DNA ENCODING SAME  
; NUMBER OF SEQUENCES: 15  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/389,722  
; FILING DATE: 04-AUG-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 369,424  
; FILING DATE: 21-JUN-1989  
; APPLICATION NUMBER: 351,117  
; FILING DATE: 12-MAY-1989  
; SEQ ID NO: 10:  
; LENGTH: 13  
5332671-10

Query Match 41.7%; Score 35; DB 5; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SFQKKTR 16  
Db 1 SFQKKTR 7

RESULT 3  
US-07-909-122-13  
; Sequence 13, Application US/07909122  
; Patent No. 5415995  
; GENERAL INFORMATION:  
; APPLICANT: SCHOOLNIK, GARY K.  
; TITLE OF INVENTION: DIAGNOSTIC PEPTIDES OF HUMAN PAPILLOMA  
; TITLE OF INVENTION: VIRUS  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/909,122  
; FILING DATE: 19920706  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENZ, WILLIAM H.  
; REGISTRATION NUMBER: 25,952  
; REFERENCE/DOCKET NUMBER: 28600-20105.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-07-909-122-13

Query Match 36.9%; Score 31; DB 1; Length 11;  
Best Local Similarity 40.0%; Pred. No. 18;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEDENQSPRS 10  
I::I::I::  
Db 1 DQDQSQTPET 10

RESULT 4  
US-08-934-915-189  
; Sequence 189, Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:  
; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, LENA  
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MASON & ASSOCIATES, P. A.  
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA

; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 3.0  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/934,915  
; FILING DATE: 22-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/949,836  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LOUISE A. FOUTCH  
; REGISTRATION NUMBER: 37,133  
; REFERENCE/DOCKET NUMBER: 1946.6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 813-538-3800  
; TELEFAX: 813-538-3820  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 189:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-934-915-189

Query Match 36.9%; Score 31; DB 2; Length 11;  
Best Local Similarity 40.0%; Pred. No. 18;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEDENQSPRS 10  
I::I::I::  
Db 1 DQDQSQTPET 10

RESULT 5  
US-08-317-522A-11  
; Sequence 11, Application US/08317522A  
; Patent No. 5599918  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/317,522A  
; FILING DATE: 04-OCT-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 9991  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-317-522A-11

Query Match 35.7%; Score 30; DB 1; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
: ||| ||  
Db 1 DOENQDPR 8

RESULT 6  
US-08-439-818A-11  
; Sequence 11, Application US/08439818A  
; Patent No. 5654145  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/439,818A  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/317,522  
FILING DATE: 04-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 1563  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 11:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-439-818A-11  
Query Match 35.7%; Score 30; DB 1; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
: ||| ||  
Db 1 DOENQDPR 8

RESULT 7  
US-08-751-965-11  
; Sequence 11, Application US/08751965  
; Patent No. 5858360  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.

TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/751,965  
FILING DATE: Herewith  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/439,818  
FILING DATE: 12-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 2252  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-751-965-11

Query Match 35.7%; Score 30; DB 2; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
: ||| ||  
Db 1 DOENQDPR 8

RESULT 8  
US-08-738-975-11  
; Sequence 11, Application US/08738975  
; Patent No. 5880267  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,975  
FILING DATE: herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/439,818

QY 2 EDENQSPR 9  
: ||| ||  
Db 1 DOENQDPR 8

; FILING DATE: 05-Dec-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 2251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-738-975-11

Query Match 35.7%; Score 30; DB 2; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
Db 1 DQENQDPR 8

RESULT 9  
US-08-728-626-11  
; Sequence 11, Application US/08728626  
; Patent No. 5910451  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/728,626  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/439,818  
; FILING DATE: 12-MAY-1995  
; APPLICATION NUMBER: US 08/317,522  
; FILING DATE: 04-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 1563  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-728-626-11

Query Match 35.7%; Score 30; DB 2; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
Db 1 DQENQDPR 8

RESULT 10  
US-08-808-599A-11  
; Sequence 11, Application US/08808599A  
; Patent No. 6111089  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin, Trophinin-Assisting  
; TITLE OF INVENTION: Proteins and Methods to Inhibit Implantation  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,599A  
; FILING DATE: 28-FEB-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/317,522  
; FILING DATE: 04-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/439,818  
; FILING DATE: 12-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 2256  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-808-599A-11

Query Match 35.7%; Score 30; DB 3; Length 9;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EDENQSPR 9  
Db 1 DQENQDPR 8

RESULT 11  
US-08-723-415B-16  
; Sequence 16, Application US/08723415B  
; Patent No. 5859199  
; GENERAL INFORMATION:  
; APPLICANT: Lathangue, Nicholas B.  
; APPLICANT: delaluna, Susana  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.



STREET: 1100 No. 5859199th Glebe Rd. 8th floor  
CITY: Arlington  
STATE: VA  
COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/723,415B  
FILING DATE: 30-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9610195.1  
FILING DATE: 15-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-220  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-723-415B-16

Query Match 35.7%; Score 30; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEDENQSPRS 10  
DB 2 EDEEDPSS 11

RESULT 12  
US-09-189-627A-16  
Sequence 16, Application US/09189627A  
Patent No. 6159691  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas  
APPLICANT: de la Luna, Susana  
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF  
FILE REFERENCE: 620-54  
CURRENT APPLICATION NUMBER: US/09/189,627A  
CURRENT FILING DATE: 1998-11-10  
PRIOR APPLICATION NUMBER: 08/723,415  
PRIOR FILING DATE: 1996-09-30  
PRIOR APPLICATION NUMBER: GB 9610195  
PRIOR FILING DATE: 1996-05-15  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: DP-3 peptide to  
OTHER INFORMATION: raise antiserum  
US-09-189-627A-16

Query Match 35.7%; Score 30; DB 3; Length 13;  
Best Local Similarity 50.0%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEDENQSPRS 10  
DB 2 EDEEDPSS 11

RESULT 13  
US-09-101-146-14  
Sequence 14, Application US/09101146  
Patent No. 6124125  
GENERAL INFORMATION:  
APPLICANT: Dartmouth College, St. Vincents Institute of  
APPLICANT: Medical Research, Kemp et al.  
TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jane Massey Licata, Esq.  
STREET: 66 E. Main Street  
CITY: Marlton  
STATE: NJ  
COUNTRY: USA  
ZIP: 08053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
COMPUTER: IBM PC  
OPERATING SYSTEM: WINDOWS 95  
SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/101,146  
FILING DATE: October 7, 1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PN7450  
FILING DATE: 8 JAN 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane Massey Licata  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: DC-0050  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (856) 810-1515  
TELEFAX: (856) 810-1454  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16  
TYPE: Amino acid  
TOPOLOGY: Linear  
US-09-101-146-14

Query Match 35.7%; Score 30; DB 3; Length 16;  
Best Local Similarity 54.5%; Pred. No. 40;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 DEDENQSPRSFOK 13  
DB 1 DEKESRRLFQQ 11

RESULT 14  
5240706-13  
Patent No. 5240706  
APPLICANT: FAULTS, DARYL  
TITLE OF INVENTION: INTRANASSAL ADMINISTRATION OF MYCOPLASMA  
HYOPNEUMONIAE ANTIGEN  
NUMBER OF SEQUENCES: 22  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/334,586  
FILING DATE: 07-APR-1989  
SEQ ID NO: 13  
LENGTH: 11  
5240706-13

Query Match 34.5%; Score 29; DB 5; Length 11;  
Best Local Similarity 83.3%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DENQSP 8  
| | | | |  
Db 6 DENQKP 11

RESULT 15  
US-08-677-298-16  
; Sequence 16, Application US/08677298  
; Patent No. 5863729  
; GENERAL INFORMATION:  
; APPLICANT: Pivnicka-Worms, Helen  
; TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAP-1  
; TITLE OF INVENTION: KINASE  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/677,298  
; FILING DATE: 09-JUL-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Caruthers, Jennie M.  
; REGISTRATION NUMBER: 34,464  
; REFERENCE/DOCKET NUMBER: 9-96  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: internal  
US-08-677-298-16

Query Match 33.3%; Score 28; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEDEQSPRS 10  
| | | | |  
Db 4 DENKEAKPRS 13

Search completed: January 17, 2001, 13:37:03  
Job time: 123 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:37:10 ; Search time 14.29 Seconds  
(without alignments)  
23.928 Million cell updates/sec

Title: US-08-765-837-3  
Perfect score: 55  
Sequence: 1 TDGSFTOPPLY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 74110

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

A\_Geneseq\_36:\*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	10	17	R90947
2	28	50.9	10	3	P20399
3	28	50.9	10	4	P30039
4	27	49.1	8	20	Y45512
5	27	49.1	8	20	W92985
6	27	49.1	10	15	Y38373
7	27	49.1	10	15	R49207
8	26	47.3	10	20	Y41894
9	24	43.6	10	15	R61602
10	23	41.8	10	17	R89241
11	23	41.8	10	20	Y48114
12	22	40.0	6	20	Y55293

13	22	40.0	6	21	Y87039
14	22	40.0	7	19	W70510
15	22	40.0	9	21	Y56067
16	22	40.0	10	15	R61553
17	22	40.0	10	17	R93565
18	22	40.0	10	18	W16026
19	22	40.0	10	19	W68664
20	22	40.0	10	19	W58143
21	22	40.0	10	20	Y14118
22	22	40.0	10	20	Y09787
23	21	38.2	6	20	Y55114
24	21	38.2	6	21	Y86860
25	21	38.2	7	12	R14205
26	21	38.2	8	17	R90387
27	21	38.2	8	20	Y41936
28	21	38.2	9	20	Y28543
29	21	38.2	10	15	R47939
30	21	38.2	10	18	W15410
31	20	36.4	6	14	R41655
32	20	36.4	6	19	W75346
33	20	36.4	6	19	W75278
34	20	36.4	6	20	Y55291
35	20	36.4	6	21	Y87037
36	20	36.4	7	17	R90407
37	20	36.4	7	20	Y55134
38	20	36.4	7	21	Y86880
39	20	36.4	8	16	R83498
40	20	36.4	8	17	W94594
41	20	36.4	8	20	Y55322
42	20	36.4	8	20	W84387
43	20	36.4	8	21	Y51163
44	20	36.4	9	18	W38395
45	20	36.4	9	20	W89119

#### ALIGNMENTS

RESULT 1  
R90947  
ID R90947 standard; peptide; 10 AA.  
XX  
AC R90947;  
XX  
DT 09-SEP-1996 (first entry)  
XX  
DE Factor VIII antigenic peptide corresp. to residues Thr1739-Tyr1748.  
XX  
KW Factor VIII; modification; inhibitor activity; binding; antibody;  
KW von Willebrand factor; immune disorder.  
XX  
OS Synthetic.  
XX  
PN WO9602572-A2.  
XX  
PD 01-FEB-1996.  
XX  
PF 14-JUL-1995; 95WO-BE00068.  
XX  
PR 14-JUL-1994; 94BE-0000666.  
XX  
PA (CRO1-) CROIX ROUGE BELGIQUE.  
XX  
PI Di Giambattista M, Laub R;  
XX  
DR Rat mACHR-6 transm  
XX  
PT Prostate specific  
XX  
PT HLA-A1 PSA antigen  
XX  
PT Rheumatoid arthrit  
XX  
PT Peptide fragment (LN clone F11 Vbeta  
XX  
XX Immunogenic peptid  
XX  
XX ATCC HB 11865 mono

*Closed 3  
my list (there)*

Human haematopoietic  
Escherichia coli g  
N-terminus of huma  
Peptide fragment (p  
Dopamine receptor  
Peptide containing  
Random peptide seq  
Interleukin-1 type  
Mutant hepatitis B  
Interleukin-1 type  
ATCC HB 11646 mono  
Human haematopoiet  
N-terminal of anti  
Antibody displacin  
Rheumatoid arthrit  
Beta-1 integrin ce  
Peptide fragment o  
Glutamate dehydrog  
Internalisation si  
Hexapeptide #1 bin  
Hexapeptide #1 bin  
ATCC HB 11865 mono  
Human haematopoiet  
Hybridoma ATCC HB-  
ATCC HB 11865 mono  
Human haematopoiet  
Zif268 mutagenised  
Human insulin anal  
HLA binding plu-1  
Finger 3 binding s  
plasmid p-DL6 pept  
Synthetic pMEL17 p  
ScFV (dig) variant

Factor VIII antigenic polypeptide fragments and epitope(s) - also  
inhibitors of factor VIII and anti-inhibitors, useful for e.g.  
preventing and treating immune disorders involving inhibition of  
factor VIII binding  
Claim 8; Page 31; 45pp; French.

XX Peptides R0945-64 are derived from the factor VIII protein, esp. from a  
 CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,  
 CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII  
 CC and derived peptides can be used to prevent the activity of inhibitors of  
 CC factor VIII binding to von Willebrand factor, esp. antibodies, thus  
 CC preventing or treating immune disorders.  
 XX  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 17; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-05;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDGSFTQPLX 10  
 Db 1 tdgstftqply 10

RESULT 2

P20399  
 ID P20399 standard; Protein; 10 AA.

XX AC P20399;  
 XX

XX 30-NOV-1992 (first entry)

XX Secretin precursor formation peptide 1.

XX Strong acid; digestive canal hormone; pancreas; gastrin; pepsin;  
 KW insulin.  
 XX  
 XX Synthetic.

Key	Location/Qualifiers
FT Modified-site 1	/note= "Boc protected"
FT Modified-site 2	/note= "Boc protected"
FT Modified-site 3	/note= "Boc protected"
FT Modified-site 5	/note= "OBUT protected"
FT Modified-site 7	/note= "BUT protected"
FT Modified-site 8	/note= "BUT protected"
FT Modified-site 9	/note= "BUT protected"
FT Modified-site	/note= "OBUT protected"

XX EP47997-A.

XX 24-MAR-1982.

XX 11-SEP-1981; 81EP-0107186.

XX 11-SEP-1980; 80JP-0125262.

XX (EISA ) EISAI KK.

XX Uchiyama M, Sato T, Yoshino H, Tsuchiya Y, Konishi M;  
 PI Tsuchiji M, Hisatake V, Koiba A;  
 XX WPI; 1982-24409E/13 (24409E).

XX Heptacosapeptide(s) - useful for high yield conversion to high  
 PT purity secretin on strong acid treatment  
 XX  
 XX Claim 4; Page 43; 47pp; English.

XX The sequence in P20398 is a precursor for the production of  
 CC secretin. The peptide sequences given in P20399-402 are peptides which

CC are useful in the production of this precursor. The precursor is  
 CC treated with strong acid in the preparation of secretin. Secretin is  
 CC one of the digestive canal hormones and is useful in promotion of  
 CC pancreatic external secretin, controlling gastrin-stimulating secretin  
 CC of the stomach acid, releasing insulin, stimulating secretin of pepsin  
 CC and decomposing fat. It is used as a pancreatic-function examining  
 CC agent and a medicine for curing duodenal ulcers etc.  
 XX  
 SQ Sequence 10 AA;

Query Match 50.9%; Score 28; DB 3; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 17;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TDGSFTQPL 9  
 Db 2 sdtftsel 10

RESULT 3

P30039  
 ID P30039 standard; peptide; 10 AA.

XX AC P30039;  
 XX

XX 04-SEP-1992 (first entry)

XX Intermediate of Secretin synthesis.

XX Porcine; digestive; hormone; pancreatic; duodenal ulcer.  
 KW Synthetic.  
 XX

Key	Location/Qualifiers
FT Modified-site 1	/label= Boc-His-Boc
FT Modified-site 2	/label= Ser-t-Bu
FT Modified-site 3	/label= Asp-O-t-Bu
FT Modified-site 5	/label= Thr-t-Bu
FT Modified-site 7	/label= Thr-t-Bu
FT Modified-site 8	/label= Ser-t-Bu
FT Modified-site 9	/label= Glu-O-t-Bu

XX JP58152852-A.

XX 10-SEP-1983.

XX 05-MAR-1982; 82JP-0034027.

XX 05-MAR-1982; 82JP-0034027.

XX (EISA ) EISAI KK.

XX WPI; 1983-791975/42.

XX Decapeptide useful as intermediate for secretin - contains  
 PT histidine, serine, aspartic acid, glycine, threonine,  
 PT phenylalanine, glutamic acid and leucine  
 XX

XX Claim 1; Page 1; 13pp; Japanese.

XX The peptide is an industrial intermediate for the production of  
 CC secretin. Secretin is a digestive tract hormone. It displays  
 CC pancreatic exocrinogenic, gastrin stimulating, gastric acid  
 CC secretion inhibiting, insulin releasing, pepsin secretion  
 CC promoting and adipsolytic action. It is used as a reagent for the

CC study of pancreatic function and as a remedy for duodenal ulcers.  
 CC See also P30038.

XX  
 SQ Sequence 10 AA;

Query Match 50.9%; Score 28; DB 4; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 17;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TDGSEFTQPL 9  
 :||:| |  
 Db 2 sdgtftsel 10

RESULT 4  
 Y45512  
 ID Y45512 standard; Peptide; 8 AA.

XX  
 AC Y45512;

XX  
 DT 01-DEC-1999 (first entry)

XX  
 DE Immunogenic peptide having a human leukocyte antigen binding motif #123.

XX  
 KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
 KW immune response; T cell activation; major histocompatibility complex;  
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
 KW vaccine; immunisation.

XX  
 OS Synthetic.  
 OS Homo sapiens.

XX  
 PN WO9945954-A1.

XX  
 PD 16-SEP-1999.

XX  
 PF 13-MAR-1998; 98WO-US05039.

XX  
 PR 13-MAR-1998; 98WO-US05039.

XX  
 PA (EPIM-) EPIMMUNE INC.

XX  
 PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

XX  
 DR WPI; 1999-551214/46.

XX  
 PT New immunogenic peptides with HLA binding motif, useful in treatment  
 PT and diagnosis of cancers and viral diseases

XX  
 PS Claim 1; Page 31; 150pp; English.

XX  
 CC Y45390 to Y48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful

CC therapeutically and for immunisation as above.

XX  
 SQ Sequence 8 AA;

Query Match 49.1%; Score 27; DB 20; Length 8;  
 Best Local Similarity 71.4%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 SFTQPLY 10  
 |||||  
 Db 2 stphply 8

RESULT 5  
 W92985

ID W92985 standard; Protein; 8 AA.

XX  
 AC W92985;

XX  
 DT 14-MAY-1999 (first entry)

XX  
 DE Rat mACHR-6 transmembrane domain #1.

XX  
 KW mACHR-6; muscarinic acetylcholine receptor 6; disorder; secretion;  
 KW acetylcholine responsive cell; phosphatidylinositol turn-over;  
 KW smooth muscle cell contraction; nervous system disorder; glandular;  
 KW schizo-effective disorder; affective disorder; sleep disorder;  
 KW movement disorder; eating disorder; drinking disorder; rat.

XX  
 OS Rattus sp.

XX  
 PN US5882893-A.

XX  
 PD 16-MAR-1999.

XX  
 PF 04-DEC-1997; 97US-0985090.

XX  
 PR 04-DEC-1997; 97US-0985090.

XX  
 PA (MILL-) MILLENNIUM PHARM INC.

XX  
 PI Goodearl AD;

XX  
 DR WPI; 1999-214063/18.

XX  
 PT Nucleic acids encoding muscarinic acetylcholine receptor 6 - useful  
 PT for modulating the effects of acetylcholine on acetylcholine  
 PT responsive cells

XX  
 PS Example 5; Column 81-82; 59pp; English.

XX  
 CC This invention describes the isolation of a novel human muscarinic  
 CC acetylcholine receptor 6 (mACHR-6), capable of modulating the effects  
 CC of acetylcholine on acetylcholine responsive cells. mACHR-6 cDNAs and  
 CC polypeptides may be used to detect naturally occurring mutations of the  
 CC mACHR-6 gene and determine if a subject with the mutated gene is at risk  
 CC of (or is predisposed to have) a mACHR-6 related disorder, modulate cell  
 CC activity mediated by mACHR-6 (e.g. biological processes mediated by  
 CC phosphatidylinositol turn-over and signalling), secretion of a molecule  
 CC (e.g. a neurotransmitter or a glandular enzyme), or contraction of a  
 CC smooth muscle cell, treat disorders mediated by abnormal mACHR-6 activity  
 CC e.g. nervous system disorders (e.g. amnesia, apraxia, agnosia, amnesic  
 CC dysnomia, amnesic spatial disorientation, Klüver-Bucy syndrome,  
 CC Alzheimer's related memory loss and learning disability, visual  
 CC hallucinations, perceptual disturbances, and Lewy body dementia  
 CC associated delirium), schizo-effective disorders (e.g. schizophrenia  
 CC with mood swings, and depressive illness), affective disorders, sleep  
 CC disorders (e.g. REM sleep abnormalities, paradoxical sleep abnormalities,  
 CC sleep-wakfulness, and body temperature or respiratory depression  
 CC abnormalities during sleep), pain generating mechanism disorders (e.g.  
 CC related to irritable bowel syndrome (IBS), or chest pain), movement  
 CC disorders (e.g. related to Parkinson's disease), eating disorders (e.g.

CC insulin hypersecretion related obesity), drinking disorders (e.g.  
 CC diabetic polydipsia), smooth muscle related disorders (e.g. IBS,  
 CC diverticular disease, urinary incontinence, oesophageal achalasia, and  
 CC chronic obstructive airways disease), cardiac disorders (e.g. pathologic  
 CC bradycardia or tachycardia, arrhythmia, flutter and fibrillation), and  
 CC glandular disorders (e.g. xerostomia and diabetes mellitus).  
 XX  
 SQ Sequence 8 AA;

Query Match 49.1%; Score 27; DB 20; Length 8;  
 Best Local Similarity 62.5%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GSFTQPLY 10  
 DB 1 gafcipy 8

RESULT 6  
 Y38373  
 ID Y38373 standard; Peptide; 10 AA.  
 XX

AC Y38373;

DT 29-SEP-1999 (first entry)

DE Prostate specific antigen-derived HLA-binding peptide.

XX Immunogen: HLA; human leukocyte antigen; binding motif; antiviral;  
 KW MHC; major histocompatibility complex; viral infection; anticancer;  
 KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.  
 XX

OS Homo sapiens.

XX WO9403205-A1.

PN 17-FEB-1994.

XX 06-AUG-1993; 93WO-US07421.

XX 05-MAR-1993; 93US-0027746.

PR 07-AUG-1992; 92US-0926666.

XX (CYTE-) CYTEL CORP.

XX Cellis E, Grey HM, Kubo RT, Sette A;

XX WPI; 1994-065403/08.

XX Peptide which specifically binds selected MHC allele - used to  
 PT induce an immune response for treatment or prevention of viral  
 PT infection or cancer, or for diagnosis  
 XX

PS Disclosure; Page 115; 150pp; English.

XX The sequence is a specific example of a group of new immunogenic  
 CC peptides having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding  
 CC motif. For example, the peptides having an HLA-A3.2 binding motif  
 CC each have 9-10 residues and contain, from the N-terminus to the  
 CC C-terminus, (a) a first conserved residue selected from L, M, I,  
 CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of  
 CC K, R, Y, H or F, where the first and second conserved residues are  
 CC separated by 6-7 residues. The peptides are capable of binding  
 CC selected MHC molecules and inducing an immune response. They can be  
 CC used to treat and/or prevent viral infection and cancer, e.g. prostate  
 CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce  
 CC antibodies for use as diagnostic or therapeutic agents. The peptides  
 CC can also be used as diagnostic agents.  
 XX

SQ Sequence 10 AA;

Query Match 49.1%; Score 27; DB 15; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 27;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 SFTQPLY 10  
 DB 4 sfphply 10

RESULT 7

R49207  
 ID R49207 standard; Protein; 10 AA.

XX R49207;

XX 31-AUG-1994 (first entry)

XX HLA-A1 PSA antigen peptide fragment 88-97.

XX Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;  
 KW immune response; viral infection; cancer; prostate cancer; lymphoma;  
 KW hepatitis; AIDS; antibody; diagnosis; prostate specific antigen.  
 XX

OS Synthetic.

XX WO9403205-A.

PN 17-FEB-1994.

XX 06-AUG-1993; 93WO-US07421.

XX 07-AUG-1992; 92US-0926666.

PR 05-MAR-1993; 93US-0027746.

XX (CYTE-) CYTEL CORP.

XX Cellis E, Grey HM, Kubo RT, Sette A;

XX WPI; 1994-065403/08.

XX Peptide which specifically binds selected MHC allele - used to  
 PT induce an immune response for treatment or prevention of viral  
 PT infection or cancer, or for diagnosis  
 XX

PS Example 8; Page 53; 150pp; English.

XX The sequences given in R47304-33 and R49201-44 are immunogenic  
 CC peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.  
 CC These peptides may be used in the composition of the invention.  
 CC These peptides are capable of binding selected MHC molecules and  
 CC inducing an immune response. They can be used to treat and/or  
 CC prevent viral infection and cancer, eg. prostate cancer, lymphoma,  
 CC hepatitis or AIDS. They can also be used to produce antibodies for  
 CC use as diagnostic or therapeutic agents. The peptides can also be  
 CC used as diagnostic agents.  
 XX

SQ Sequence 10 AA;

Query Match 49.1%; Score 27; DB 15; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 27;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 SFTQPLY 10  
 DB 4 sfphply 10

RESULT 8

Y41894

ID Y41894 standard; Peptide; 10 AA.

XX Y41894;

XX 09-DEC-1999 (first entry)  
 DT Rheumatoid arthritis diagnostic protein isoform peptide #45.  
 DE Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;  
 XX rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;  
 KW rheumatoid arthritis diagnostic protein isoform; screening;  
 KW expression reference protein isoform; prognosis.  
 XX Homo sapiens.  
 OS WO9947925-A2.  
 XX 23-SEP-1999.  
 PN 15-MAR-1999; 99WO-GB00763.  
 XX 13-MAR-1998; 98GB-0005477.  
 XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 PA Parekh RB, Patel TP, Townsend RR;  
 PI WPI; 1999-571871/48.  
 DR Diagnosis of human rheumatoid arthritis by two-dimensional  
 PT electrophoresis -  
 PT Disclosure; Page 18; 157pp; English.  
 PS A method has been developed for the diagnosis of human rheumatoid  
 CC arthritis (RA) using two-dimensional electrophoresis to generate a  
 CC two-dimensional array of features. The method can be used for screening,  
 CC diagnosis and prognosis of RA in a subject or for monitoring the effect  
 CC of an anti-RA drug or therapy administered to a subject. The method  
 CC comprises: (a) analysing a sample of serum or plasma and optionally  
 CC synovial fluid by two-dimensional electrophoresis, to generate a two-  
 CC dimensional array of features; (b) identifying at least one chosen  
 CC feature whose relative abundance correlates with the presence or absence  
 CC of RA; and (c) comparing the abundance of each chosen feature in the  
 CC sample with the abundance of that chosen feature in serum or plasma from  
 CC one or more persons without RA, where the relative abundance of the  
 CC chosen feature or features in the sample indicates the presence or  
 CC absence of RA in the subject. The method can also be used in clinical  
 CC studies for testing drugs for therapy of RA, for purification of RA-  
 CC diagnostic protein isoforms (RPIs), and for production of antibodies to  
 CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify  
 CC compounds that promote or inhibit their activity, which are then used as  
 CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy  
 CC protocols. Y41844 to Y42100 represent RPI peptides Y42101 to Y42103  
 CC represent expression reference protein isoform peptides and 225066 to  
 CC 225068 represent degenerate probes for RPIs, which are all used in  
 CC the exemplification of the present invention.  
 XX Sequence 10 AA;  
 SQ

Query Match 47.3%; Score 26; DB 20; Length 10;  
 Best Local Similarity 83.3%; Pred. No. 42;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDGSFT 6  
 II III  
 Db 3 tdaaft 8

RESULT 9  
 R61602  
 ID R61602 standard; peptide; 10 AA.  
 XX R61602;  
 AC  
 KW

DT 12-MAY-1995 (first entry)  
 XX Peptide fragment (1.0796) of HBV binds HLA-A2.1.  
 DE antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; HIV;  
 XX HIVI; plasma specific antigen; hepatitis B virus; Epstein Barr;  
 KW human immunodeficiency virus; human papilloma virus; p53; C-ERB2;  
 KW MAGP-1; melanoma antigen-1; core antigen; surface antigen;  
 KW pharmaceutical composition; in vivo; ex vivo; therapeutic;  
 KW diagnostic; MHC class I molecule; major histocompatibility complex;  
 KW HLA-A2.1; 9mer; 10mer; anchor; human leukocyte antigen.  
 XX Hepatitis B virus strain adw.  
 OS WO9420127-A.  
 XX 15-SEP-1994.  
 PN 04-MAR-1994; 94WO-US02353.  
 XX 05-MAR-1993; 93US-0027146.  
 PR 04-JUN-1993; 93US-0073205.  
 PR 29-NOV-1993; 93US-0159184.  
 XX (CYTE-) CYTEL CORP.  
 PA Grey HM, Kast WM, Sette A, Sidney J;  
 PI WPI; 1994-302678/37.  
 DR Immunogenic peptide(s) having an HLA-A2.1 binding motif - used  
 XX for treatment or prophylaxis of cancer, virus infection or  
 PT autoimmune diseases.  
 PT Example 5; Page 111; 138pp; English.  
 PS R59496-R61666 are immunogenic 10mer peptides that contain a HLA-A2.1  
 CC binding motif. These peptides bind HLA-A2.1 and have a binding  
 CC affinity of at least 1% as compared to a reference peptide (R71293).  
 CC R61602 has an IC50 of 0.0002 and the sequence occurs at position  
 CC 1317 in the HBV POL protein. The peptides of the invention can induce  
 CC cytotoxic T lymphocytes which can react with target cells. They can  
 CC be used for the treatment or prophylaxis of cancer, eg. prostate  
 CC cancer or lymphoma, etc.  
 XX Sequence 10 AA;  
 SQ

Query Match 43.6%; Score 24; DB 15; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 1e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GSFTQPL 9  
 I: I I  
 Db 4 gtfvspl 10

RESULT 10  
 R89241  
 ID R89241 standard; peptide; 10 AA.  
 XX R89241;  
 AC  
 XX 05-SEP-1996 (first entry)  
 DT LN clone F11 Vbeta8-CDR3.  
 XX Polymerase chain reaction; PCR; primer; amplify; human; T cell receptor;  
 KW beta chain; TCR; myelin basic protein; BP; autoantigen; encephalitogen;  
 KW experimental autoimmune encephalomyelitis; EAE; multiple sclerosis; MS;  
 KW autoimmune disease; neurological disease; cerebrospinal fluid; therapy;  
 KW central nervous system; complementarity determining region; CDR;  
 KW T lymphocyte; optical nerve damage; anterior chamber inflammation.

XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 3 /note= "encoded by GGC"  
 XX XX  
 XX PN W09601329-A1.  
 XX XX  
 XX PD 18-JAN-1996.  
 XX XX  
 XX PF 26-JUN-1995; 95WO-US08086.  
 XX XX  
 XX PR 01-JUL-1994; 94US-0270634.  
 XX XX  
 XX PA (CONN-) CONNECTIVE THERAPEUTICS INC.  
 XX XX  
 XX PI Buenafe A, Offner H, Vandenbark AA;  
 XX XX  
 XX DR WPI; 1996-087679/09.  
 XX DR N-PSDB; T10618.  
 XX XX  
 XX PT Methods for diagnosis and immune-related therapy of autoimmune  
 PT diseases - partic. multiple sclerosis, by detecting marker T cell  
 PT receptor V gene bias and treating patients with selected V beta  
 PT peptide(s)  
 XX XX  
 XX PS Example 2; Fig 6c; 62pp; English.  
 XX XX  
 CC R99215-R99251 represent clones of the Vbeta8 complementarity determining  
 CC region 3 (CDR3) of the T cell receptor beta (TCRBeta) chain. These  
 CC sequences were isolated from cerebrospinal fluid (CSF), spinal cord (SC)  
 CC and lymph nodes (LN) of clones of Lewis rats with experimental autoimmune  
 CC encephalomyelitis (EAE). By detecting the presence of a marker TCR V  
 CC gene bias in a body fluid which encapsulates all or part of the target  
 CC organ, an autoimmune disease (such as a neurological disease) in a human  
 CC can be identified. This method can also be carried out to detect the  
 CC presence of a biased motif common to T cell receptors specific for the  
 CC pathogenic antigen in a non-target tissue or organ. By analysing the  
 CC Vbeta gene repertoire of CSF, and determining the presence of a Vbeta  
 CC gene bias, an immune-related disease that targets the central nervous  
 CC system can be diagnosed. Therapeutic Vbeta peptide sequences can be  
 CC selected to use as treatment of a disease or condition. The selection is  
 CC carried out by identifying a Vbeta gene bias in a body fluid that is not  
 CC the target tissue or organ of the disease, and selecting an immunogenic  
 CC peptide corresponding to the Vbeta gene bias. Multiple sclerosis (MS)  
 CC can be treated by identifying the CDR2 of a V gene peptide on the surface  
 CC of a T lymphocyte in the CSF of a patient and administering a peptide  
 CC corresponding to this region. These methods can also be used for the  
 CC diagnosis and immune-related therapy of optical nerve damage and anterior  
 CC chamber inflammation as well as other human neurological diseases.  
 XX XX  
 SQ Sequence 10 AA;

Query Match 41.8%; Score 23; DB 17; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 TQPLY 10  
 I I I I  
 Db 5 tsply 9

RESULT 11  
 Y48114  
 ID Y48114 standard; Peptide; 10 AA.  
 XX XX  
 AC Y48114;  
 XX XX  
 DT Q1-DEC-1999 (first entry)  
 XX XX  
 DE Immunogenic peptide having a human leukocyte antigen binding motif #2725.

XX Human leukocyte antigen: binding; immunogenic; glycoprotein; MHC; HLA;  
 KW immune response; T cell activation; major histocompatibility complex;  
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
 KW vaccine; immunisation.  
 XX XX  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09945954-A1.  
 XX XX  
 XX PD 16-SEP-1999.  
 XX XX  
 XX PF 13-MAR-1998; 98WO-US050509.  
 XX XX  
 XX PR 13-MAR-1998; 98WO-US050509.  
 XX XX  
 XX PA (EPIM-) EPIMMUNE INC.  
 XX XX  
 XX PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;  
 XX XX  
 XX DR WPI; 1999-551214/46.  
 XX XX  
 XX PT New immunogenic peptides with HLA binding motif, useful in treatment  
 PT and diagnosis of cancers and viral diseases  
 XX XX  
 XX PS Claim 1; Page 137; 150pp; English.  
 XX XX  
 CC Y45390 to Y48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful  
 CC therapeutically and for immunisation as above.  
 XX XX  
 SQ Sequence 10 AA;

Query Match 41.8%; Score 23; DB 20; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 FTQPLY 10  
 I I I I  
 Db 1 fphply 6

RESULT 12  
 Y55293  
 ID Y55293 standard; peptide; 6 AA.  
 XX XX  
 AC Y55293;  
 XX XX  
 DT 07-JAN-2000 (first entry)  
 XX XX  
 DE ATCC HB 11885 monoclonal antibody 9079 releasing peptide SEQ ID NO:187.  
 XX XX



KW Antibody releasing peptide; CD34; hybridoma; binding; antigen;  
 KW cell surface antigen; identification; haematopoietic stem cell;  
 KW tumour; cancer; immune system; therapy; displacement.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN US5968753-A.  
 XX  
 XX 19-OCT-1999.  
 PD  
 XX  
 XX 07-JUN-1995; 95US-0482228.  
 PF  
 XX  
 XX 14-JUN-1994; 94US-0259427.  
 PR  
 XX  
 XX (NEXE-) NEXELL THERAPEUTICS INC.  
 PA  
 XX  
 PI Guillermo R, Helgerson SL, Deans RJ, Tseng-Law J, Kobori JA;  
 PI Al-Abdaly FA;  
 XX  
 DR WPI; 1999-590399/50.  
 XX  
 XX Short peptides useful for displacing antibodies from cell surface  
 PT antigens.  
 PT  
 XX  
 PS Example 9; Column 33; 81pp; English.  
 XX  
 XX The present invention describes peptides of 4-17 amino acids which  
 CC displace either the anti-CD34 monoclonal antibody designated 561, the  
 CC anti-CD34 mouse monoclonal antibody produced by the hybridoma ATCC  
 CC HB-11646 (designated 9069), the anti-CD34 antibody produced by  
 CC hybridoma ATCC HB-11885 (9079), or the anti-human breast cancer  
 CC antibody produced by hybridoma ATCC HB-11884 (9187), from a cell  
 CC surface antigen on a target cell. The peptides are useful for  
 CC displacing antibodies bound to cell surfaces to release cells that  
 CC have been positively selected by antibody-mediated binding to beads  
 CC or other solid support. Y55107 to Y55319 represent peptides used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 6 AA;  
  
 Query Match 40.0%; Score 22; DB 20; Length 6;  
 Best Local Similarity 60.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 5 FTQPL 9  
 Db :|||:  
 2 ytcpi 6  
  
 RESULT 13  
 Y87039  
 ID Y87039 standard; Peptide; 6 AA.  
 XX  
 AC Y87039;  
 XX  
 XX 09-MAY-2000 (first entry)  
 DT  
 XX Human haematopoietic CD34+ cell binding peptide SEQ ID #187.  
 DE  
 XX Human; haematopoietic CD34+ cell; binding peptide; monoclonal antibody;  
 KW non-enzymatic cell selection method; haematopoietic stem cell;  
 KW haematopoietic progenitor cell; antibody 561; breast cancer cell;  
 KW antibody 9187; cell surface determinant; diagnostic cell based assay.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US6017719-A.  
 PN  
 XX 25-JAN-2000.  
 PD  
 XX 07-JUN-1995; 95US-0482528.  
 PF

XX  
 PR 14-JUN-1994; 94US-0259427.  
 XX  
 PA (NEXE-) NEXELL THERAPEUTICS INC.  
 XX  
 PI Guillermo R, Helgerson SL, Deans RJ, Tseng-Law J, Kobori JA;  
 PI Al-Abdaly FA;  
 XX  
 DR WPI; 2000-136676/12.  
 XX  
 XX Non-enzymatic method for the positive selection of target cells from a  
 PT heterogeneous cell suspension, useful for selecting human breast cancer  
 PT cells from a patient's blood or bone marrow -  
 XX  
 XX Example 9; Column 38; 82pp; English.  
 PS  
 XX This sequence represents a human haematopoietic CD34+ cell binding  
 CC peptide, and was used to test the method of the invention. The method is:  
 CC a non-enzymatic method for the positive selection of one or more target  
 CC cells from a heterogeneous cell suspension, by using specific peptides  
 CC which effect the displacement and release of a specific target cell  
 CC from a specific monoclonal antibody. The method is useful for positive  
 CC selection and specific release of target human haematopoietic  
 CC stem/progenitor cells bound by the monoclonal anti-CD34 antibodies and  
 CC the antibody 561. The method is also useful for positive selection and  
 CC specific release of target human breast cancer cells, bound by the  
 CC monoclonal anti-breast cancer antibody 9187, from a patient's blood or  
 CC bone marrow. Identification of peptide epitopes for antibodies which  
 CC recognise cell surface determinants also allows construction of  
 CC diagnostic cell based assays. The peptide mediated release is enzyme free,  
 CC and thus leaves the cell surface proteins intact. Moreover, peptide  
 CC mediated release leaves the target cell free of bound antibody or  
 CC antibody fragments. The method also produces a high yield of functional  
 CC target cells and is relatively inexpensive to carry out.  
 XX  
 SQ Sequence 6 AA;  
  
 Query Match 40.0%; Score 22; DB 21; Length 6;  
 Best Local Similarity 60.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 5 FTQPL 9  
 Db :|||:  
 2 ytcpi 6  
  
 RESULT 14  
 W70510  
 ID W70510 standard; peptide; 7 AA.  
 XX  
 AC W70510;  
 XX  
 DT 29-DEC-1998 (first entry)  
 XX  
 XX Escherichia coli glucose dehydrogenase mutant region peptide sequence 2.  
 DE  
 XX Escherichia coli glucose dehydrogenase; GDH; PQQ; PQQDH; glucose;  
 KW pyroloquinoline quinone coenzyme.  
 KW  
 KW Escherichia coli.  
 XX  
 OS Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 4  
 FT /note= "Changed from His in wild-type to Asp in  
 FT mutant"  
 XX  
 XX JP10243786-A.  
 PN  
 XX 14-SEP-1998.  
 PD  
 XX 03-MAR-1997; 97JP-0061727.  
 PF

```

XX 03-MAR-1997; 97JP-0061727.
XX (HAYA/) HAYADE H.
XX WPT: 1998-549782/47.
XX Modified glucose dehydrogenase - allows sensitive and selective
XX quantitative determination of glucose
XX PT
XX Disclosure; Page -: 5pp; Japanese.
XX
XX The present sequence represents a modified region (residues 772-778)
XX of the Escherichia coli glucose dehydrogenase (GDH) comprising of a
XX single residue mutation H775D, i.e. where residue 4 (representing
XX residue 775 of GDH) is changed from His to Asp. The invention
XX provides modified E. coli GDH enzymes, which are modified at residue
XX 775, with their pyrroloquinoline quinone (PQQ) coenzyme. A GDH enzyme
XX with its PQQ coenzyme is known as PQQDH. The invention also provides
XX the gene encoding the modified PQQDH enzyme. The modified PQQDH
XX enzymes are claimed to allow highly sensitive and selective quantitative
XX determination of glucose.
XX Note: The present sequence was not provided by the inventors but was
XX made up using the wild-type region, representing residues 772-778 of
XX GDH, and other relevant information provided in the specification.
XX
SQ Sequence 7 AA;

Query Match 40.0%; Score 22; DB 19; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DGSF 5
Db 4 dgsf 7

RESULT 15
Y56067
ID Y56067 standard; peptide; 9 AA.
XX
XX Y56067;
XX
XX 11-APR-2000 (first entry)
XX
XX N-terminus of human Vbeta17 chain HLA-A0201 restricted TCR.
XX
XX Fusion protein; T-cell receptor; c-jun; leucine zipper motif; tumour;
XX influenza virus matrix protein epitope; major histocompatibility complex;
XX immune system; PCR primer; mutagenesis.
XX
XX Homo sapiens.
XX
XX WO9960119-A2.
XX
XX 25-NOV-1999.
XX
XX 19-MAY-1999; 99WO-GB01583.
XX
XX 19-MAY-1998; 98GB-0010759.
XX 29-SEP-1998; 98GB-0021129.
XX
XX (AVID-) AVIDEX LTD.
XX
XX Jakobsen BK, Boulter JM;
XX
XX WPI; 2000-072439/06.
XX N-PSDB; 247257.
XX
XX Synthetic multivalent receptor complexes used to detect Major
XX Histocompatibility Complex-peptide complexes, and delivering
XX therapeutic agents to target cells

```

Example 10; Fig 20B; 155pp; English.

The invention relates to a synthetic multivalent complex for binding to a major histocompatibility complex (MHC)-peptide complex. The multivalent complex comprises a T-cell receptor (TCR) protein, preferably fused to a leucine zipper sequence from c-jun or c-fos for stabilising the TCR complex. Sequences Y56066-Y56075 represent the N-termini of alpha and beta chain genes from a range of HLA restricted TCR proteins. This sequence represents the N-terminus of the human Vbeta17 chain from an influenza matrix peptide HLA-A0201 restricted TCR, encoded by clone JM22. The multivalent TCR complexes are useful for detecting MHC-peptide complexes and delivery of therapeutic agents to target cells, in vivo. The specificity of the TCRs enables the localization of the liposome-contained drugs to the desired target site such as a tumour or virus-infected cell. This would be useful in many situations and in particular against tumours because not all cells in the tumour present antigens and therefore are not all detected by the immune system.

Sequence 9 AA;

Query Match 40.0%; Score 22; DB 21; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2.1e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DGSFTQ 7  
Db 3 dggitq 8

Search completed: January 17, 2001, 13:39:47  
Job time: 157 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:38:41 ; Search time 13.63 seconds  
(without alignments)  
49.817 Million cell updates/sec

Title: US-08-765-837-3  
Perfect score: 55  
Sequence: 1 TDGSFTQPLY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues  
Total number of hits satisfying chosen parameters: 1097

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	32.7	7	2 A58718	carnocin UI49 - Ca
2	18	32.7	10	2 C61440	polygalacturonase
3	18	32.7	10	2 PH0948	T-cell receptor be
4	17	30.9	10	2 PH1344	Ig heavy chain DJ
5	16	29.1	6	2 I51317	bHLH transcription
6	16	29.1	8	2 A23967	leucoperoxin - M
7	16	29.1	9	2 A37027	macrophage chemota
8	16	29.1	9	2 PD0443	3-oxoacid CoA-tran
9	16	29.1	10	2 S38305	lectin GNL2 alpha
10	16	29.1	10	2 PT0309	Ig heavy chain CRD
11	15	27.3	5	2 PT0644	T-cell receptor be
12	15	27.3	7	2 E33932	Ig mu chain D regi
13	15	27.3	8	2 PT0030	inulinase (EC 3.2.
14	15	27.3	9	2 S36898	ribosomal protein
15	15	27.3	10	2 S38844	alpha-2-macroglobu
16	15	27.3	10	2 C39191	hypothetical prote
17	15	27.3	10	2 A61007	hementin (EC 3.4.
18	14	25.5	6	2 B61512	variant surface gl
19	14	25.5	6	2 PT0280	Ig heavy chain CRD
20	14	25.5	7	2 PT0611	T-cell receptor be
21	14	25.5	8	2 S21288	lectin - potato (f
22	14	25.5	9	2 S55696	phosphoenolpyruvat
23	14	25.5	9	2 JN0027	[Phe-6]-mosact - s
24	14	25.5	9	2 G41946	T-cell receptor ga
25	14	25.5	10	2 PC2171	triacylglycerol li
26	14	25.5	10	2 E61512	variant surface gl
27	14	25.5	10	2 A61622	vitellogenin, 190k
28	14	25.5	10	2 PT0243	Ig heavy chain CRD
29	14	25.5	10	2 A90345	angiotensin precu

30	14	25.5	10	2 PH0925	T-cell receptor be
31	13	23.6	4	2 A02147	phagocytosis-stimu
32	13	23.6	4	2 PT0271	Ig heavy chain CRD
33	13	23.6	5	2 PT0699	T-cell receptor be
34	13	23.6	6	2 PT0657	T-cell receptor be
35	13	23.6	6	2 PT0720	T-cell receptor be
36	13	23.6	6	2 PT0589	T-cell receptor be
37	13	23.6	6	2 A43129	neuropeptide GNPFR
38	13	23.6	7	2 PC1316	large granule l3 c
39	13	23.6	7	2 PT0523	T-cell receptor be
40	13	23.6	7	2 PT0689	T-cell receptor be
41	13	23.6	7	2 A39690	neural cell adhesi
42	13	23.6	7	4 A58725	virotoxin - destro
43	13	23.6	8	2 A32523	peptidyl-dipeptida
44	13	23.6	8	2 S20162	leghemoglobin III
45	13	23.6	8	2 XGHUEU	urine glycopeptide

ALIGNMENTS

RESULT 1  
A58718  
carnocin UI49 - Carnobacterium sp. (fragment)  
C:Species: Carnobacterium sp.  
C:Date: 23-Jan-1998 #sequence\_revision 30-Jan-1998 #text\_change 30-Jan-1998  
C:Accession: A58718  
R:Stoffels, G.; Nissen-Meyer, J.; Gudmundsdottir, A.; Sletten, K.; Holo, H.; Nes, I.F  
Appl. Environ. Microbiol. 58, 1417-1422, 1992  
A:Title: Purification and characterization of a new bacteriocin isolated from a Carno  
A:Reference number: A58718; MUID:92321768  
A:Accession: A58718  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <STO>  
C:Keywords: antibiotic; lanthionine

Query Match 32.7%; Score 18; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GSFTOP 8  
Db 1 GSEIQP 6

RESULT 2  
C61440  
polygalacturonase (EC 3.2.1.15) II b - Aspergillus sp. (fragment)  
C:Species: Aspergillus sp.  
C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 06-Dec-1996  
C:Accession: C61440  
R:Stratilova, E.; Markovic, O.; Skrovinova, D.; Rexova-Benkova, L.; Jorncvalli, H.  
J. Protein Chem. 12, 15-22, 1993  
A:Title: Pectinase Aspergillus sp. polygalacturonase: multiplicity, divergence, and s  
A:Reference number: A61440; MUID:93151962  
A:Accession: C61440  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <STR>  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 32.7%; Score 18; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GSFT 6  
Db 3 GAFT 6

## RESULT 3

PH0948  
T-cell receptor beta chain V-D-J region - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 19-Oct-1995 #text\_change 30-May-1997  
C:Accession: PH0948; PH0897; PH0909; PH0899  
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandembark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic  
A:Reference number: PH0891; MUID:92078857  
A:Accession: PH0948  
A:Molecule type: mRNA  
A:Residues: 1-10 <GO1>  
A:Experimental source: myelin basic protein fragment-reactive T-cell, recovered from exp  
A:Accession: PH0897  
A:Molecule type: mRNA  
A:Residues: 1-10 <GO2>  
A:Experimental source: myelin basic protein-immunized T-cell, clones 3, 6-2, 14, hybrid  
A:Accession: PH0909  
A:Molecule type: mRNA  
A:Residues: 1-10 <GO3>  
A:Accession: PH0899  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-10 <GO4>  
A:Experimental source: clone 14  
C:Keywords: T-cell receptor

Query Match 32.7%; Score 18; DB 2; Length 10;  
Best Local Similarity 42.9%; Pred. No. 1.4e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TDGSFTQ 7  
: : : :  
Db 4 SDSSEFQ 10

## RESULT 4

PH1344  
Ig heavy chain DJ region (clone C100-91A) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: PH1344  
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.  
J. Exp. Med. 176, 1577-1581, 1992  
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph  
A:Reference number: PH1302; MUID:93094761  
A:Accession: PH1344  
A:Molecule type: DNA  
A:Residues: 1-10 <WAS>  
C:Keywords: heterotetramer; immunoglobulin

Query Match 30.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 42.9%; Pred. No. 2.2e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TDGSFTQ 7  
: : : :  
Db 2 SDGDWQ 8

## RESULT 5

I51317  
bHLH transcription factor inhibitor - African clawed frog (fragment)  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I51317  
R:Zhang, H.; Reynaud, S.; Kloc, M.; Etkin, L.D.; Spohr, G.  
Mech. Dev. 50, 119-130, 1995  
A:Title: Id gene activity during Xenopus embryogenesis.  
A:Reference number: I51316; MUID:95344988

A:Accession: I51317  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-6 <ZNA>  
A:Cross-references: GB:S79038; NID:gl042006; PIDN:AAD14294.1; PID:94261994  
C:Genetics:  
A:Gene: Xidib

Query Match 29.1%; Score 16; DB 2; Length 6;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 SFTQPL 8  
: : : :  
Db 1 SATEP 5

## RESULT 6

A23967  
leucopyrokinin - Madeira cockroach  
C:Species: Leucophaea maderae (Madeira cockroach)  
C:Date: 31-Mar-1988 #sequence\_revision 26-May-1994 #text\_change 11-Jul-1997  
C:Accession: A23967  
R:Nachman, R.J.; Holman, G.M.; Cook, B.J.  
Biochem. Biophys. Res. Commun. 137, 936-942, 1986  
A:Title: Active fragments and analogs of the insect neuropeptide leucopyrokinin: stru  
A:Reference number: A23967; MUID:86269041  
A:Accession: A23967  
A:Molecule type: protein  
A:Residues: 1-8 <NAC>  
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 29.1%; Score 16; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 SFTQPL 9  
: : : :  
Db 3 SFTPL 8

## RESULT 7

A37027  
macrophage chemotactic factor - human (fragment)  
N:Alternate names: T-cell hybridoma D6-18 protein  
C:Species: Homo sapiens (man)  
C:Date: 31-Jan-1992 #sequence\_revision 31-Jan-1992 #text\_change 23-Feb-1997  
C:Accession: A37027  
R:Yoshizuka, N.; Yoshimura, M.; Tsuchiya, S.; Okamoto, K.; Kobayashi, Y.; Osawa, T.  
Cell. Immunol. 123, 212-225, 1989  
A:Title: Macrophage chemotactic factor (MCF) produced by a human T cell hybridoma clo  
A:Reference number: A37027; MUID:89376581  
A:Accession: A37027  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <YOS>

Query Match 29.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DGS 4  
: : : :  
Db 6 DGS 8

## RESULT 8

PD0443

3-oxoacid CoA-transferase (EC 2.8.3.5) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 05-Feb-1999  
 C:Accession: PD0443

R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.

submitted to JIPID, August 1998

A:Description: Proteome analysis of mouse brain.

A:Reference number: PD0441

A:Contents: Striatum

A:Accession: PD0443

A:Molecule type: protein

A:Residues: 1-9 <KAW>

C:Keywords: CoA-transferase

Query Match 29.1%; Score 16; DB 2; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 2e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 FTQPL 9

I I I

Db 4 YTPDV 8

RESULT 9

S38305

lectin GNL2 alpha chain - kidney bean (fragment)

C:Species: Phaseolus vulgaris (kidney bean)

C:Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 13-Mar-1997

C:Accession: S38305

R:Kamemura, K.; Furuichi, Y.; Umekawa, H.; Takahashi, T.

Biochim. Biophys. Acta 1158, 181-188, 1993

A:Title: Purification and characterization of novel lectins from Great Northern bean, Ph

A:Reference number: S38304; MUID:94002183

A:Accession: S38305

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <KAM>

Query Match 29.1%; Score 16; DB 2; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 3.4e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDGSET 6

I I I

Db 2 TETSFS 7

RESULT 10

PT0309

Ig heavy chain CDR3 region (clone 6-94) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C:Accession: PT0309

R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and

A:Reference number: PT0222; MUID:91108337

A:Accession: PT0309

A:Molecule type: DNA

A:Residues: 1-10 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 29.1%; Score 16; DB 2; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 3.4e+03;  
 Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 DGSFTQPLY 10

I I I

Db 2 DCGNSRAAY 10

RESULT 11

PT0644

T-cell receptor beta chain V-D-J region (111-16) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PT0644

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0644

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <FEE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 27.3%; Score 15; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SFT 6

I I I

Db 3 SFT 5

RESULT 12

E33932

Ig mu chain D region (E7) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 16-Aug-1996

C:Accession: E33932

R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.

Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989

A:Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated chr

A:Reference number: A33932; MUID:89282823

A:Accession: E33932

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-7 <BAC>

A:Cross-references: GB:M27106

C:Keywords: immunoglobulin

Query Match 27.3%; Score 15; DB 2; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 OPLY 10

I I I

Db 1 RPFY 4

RESULT 13

PT0030

inulinase (EC 3.2.1.7) - Aspergillus ficuum (fragment)

N:Alternate names: inulase

C:Species: Aspergillus ficuum

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 07-May-1999

C:Accession: PT0030

R:Etalibi, M.; Baratti, J.C.

Agric. Biol. Chem. 54, 61-68, 1990

A:Title: Molecular and kinetic properties of Aspergillus ficuum inulinases.

A:Reference number: PT0030; MUID:90344234

A:Accession: PT0030

A:Molecule type: protein

A:Residues: 1-8 <ETT>

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 27.3%; Score 15; DB 2; Length 8;  
Best Local Similarity 40.0%; Pred. No. 2e+05;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 SFTQP 8  
: : :  
Db 2 NYDQP 6

## RESULT 14

S36898  
ribosomal protein S8 - Mycobacterium bovis (fragment)  
C:Species: Mycobacterium bovis  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
C:Accession: S36898  
R:Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.  
FEBS Lett. 331, 9-14, 1993  
A:Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycobacterium bovis  
A:Reference number: S36887; MUID:94009653  
A:Accession: S36898  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <OHA>  
C:Keywords: protein biosynthesis; ribosome

Query Match 27.3%; Score 15; DB 2; Length 9;  
Best Local Similarity 33.3%; Pred. No. 2e+05;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 SFTQPL 9  
: : :  
Db 1 TMTDPI 6

## RESULT 15

S33844  
alpha-2-macroglobulin - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 22-Nov-1993 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C:Accession: S33844  
R:Warburton, M.J.; Coles, B.; Dundas, S.R.; Gusterson, B.A.; O'Hare, M.J.  
Eur. J. Biochem. 214, 803-809, 1993  
A:Title: Hydrocortisone induces the synthesis of alpha(2)-macroglobulin by rat mammary epithelial cells  
A:Reference number: S33843; MUID:93307297  
A:Accession: S33844  
A:Molecule type: protein  
A:Residues: 1-10 <WAR>  
C:Superfamily: alpha-2-macroglobulin  
C:Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond

Query Match 27.3%; Score 15; DB 2; Length 10;  
Best Local Similarity 44.4%; Pred. No. 5.3e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 2 DGSFTQPLY 10  
: : :  
Db 3 DGK--EPQY 9

Search completed: January 17, 2001, 13:40:19  
Job time: 98 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: January 17, 2001, 13:39:51 ; Search time 8.5 Seconds  
(without alignments)  
37.993 Million cell updates/sec

Title: US-08-765-837-3  
Perfect score: 55  
Sequence: 1 TDGFTQPLY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 329

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	36.4	8	1 LMT2_LOCOMI	P22396 locusta mig
2	19	34.5	10	1 ANGL_BOTJA	Q10581 bothrops ja
3	18	32.7	7	1 LANC_CARUI	P36960 carnobacter
4	18	32.7	8	1 ALI6_CVDPO	P82157 cydia pomon
5	16	29.1	6	1 VP19_HSVIK	P23210 herpes simp
6	16	29.1	8	1 LCK4_LEUMA	P21143 leucophaea
7	16	29.1	8	1 LCK4_LEUMA	P13049 leucophaea
8	15	27.3	10	1 TOP1_SALTY	P40686 salmonella
9	14	25.5	9	1 MOSF_GLYJA	P19853 clypeaster
10	14	25.5	9	1 UPA3_HUMAN	P30089 homo sapien
11	14	25.5	10	1 ANGT_BOVIN	P01017 bos taurus
12	13	23.6	4	1 TUFT_HUMAN	P01858 homo sapien
13	13	23.6	6	1 FARP_MONEX	P41966 moniezia ex
14	13	23.6	7	1 MNP1_LEPDE	P42984 leptinotars
15	13	23.6	8	1 GLUR_HUMAN	P02729 homo sapien
16	13	23.6	10	1 AMPN_HELAM	P81731 helicoverpa
17	13	23.6	10	1 LPR2_LOCOMI	P41488 locusta mig
18	13	23.6	10	1 SP34_DICMU	P81545 dictyosteli
19	12	21.8	6	1 CIP1_MYTED	P13736 mytilus edu
20	12	21.8	6	1 CIP2_MYTED	P13737 mytilus edu
21	12	21.8	8	1 AKH_LIBAU	P25418 libellula a
22	12	21.8	8	1 AKH_TABAT	P14595 tabanus atr
23	12	21.8	8	1 FARI_PANRE	P41872 panagrellus
24	12	21.8	8	1 FARI3_HOMAM	P41486 homarus ane
25	12	21.8	8	1 HTE2_PERAM	P04549 periplaneta
26	12	21.8	8	1 LCK7_LEUMA	P19989 leucophaea
27	12	21.8	8	1 NP8_BOVIN	P15507 bos taurus
28	12	21.8	9	1 ALI0_CARMA	P81813 carcinus ma
29	12	21.8	9	1 CCAP_CARMA	P38556 carcinus ma
30	12	21.8	9	1 COXE_THUOB	P80975 thunnus obe
31	12	21.8	9	1 FARI9_ASCSU	P43172 ascaris suu
32	12	21.8	9	1 FARD_CALVO	P41868 calliphora
33	12	21.8	9	1 FIBB_PAPAN	P19344 papio anubi

34	12	21.8	9	1 FIBB_PAPHA	P19343 papio hamad
35	12	21.8	9	1 FIBB_THEGE	P19342 theropithec
36	12	21.8	9	1 LMT3_LOCOMI	P41489 locusta mig
37	12	21.8	10	1 ANGT_CHICK	P01018 gallus gall
38	12	21.8	10	1 GNL1_ALLMI	P37041 alligator m
39	12	21.8	10	1 HTF1_ROMMI	P18110 romalea mic
40	12	21.8	10	1 HTF2_CARMO	P11385 carausius m
41	12	21.8	10	1 HTF_TABAT	P14596 tabanus atr
42	12	21.8	10	1 PPCK_FASHE	P80525 fasciola he
43	12	21.8	10	1 SPI_HALRO	Q10997 halocynthia
44	12	21.8	10	1 UH05_RAT	P56573 rattus norv
45	12	21.8	10	1 UPA5_HUMAN	P30091 homo sapien

ALIGNMENTS

RESULT 1					
LMT2_LOCOMI					
ID LMT2_LOCOMI	STANDARD;	PRT;	8 AA.		
AC P22396;					
DT 01-AUG-1991 (Rel. 19, Created)					
DT 01-AUG-1991 (Rel. 19, Last sequence update)					
DT 01-AUG-1991 (Rel. 19, Last annotation update)					
DE LOCUSTAMYOTROPIN 2 (LOM-MT-2).					
OS Locusta migratoria (Migratory locust).					
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;					
OC Acridomorpha; Acridoidea; Acrididae; Locusta.					
RN [1]					
RP SEQUENCE.					
RC TISSUE=CORPORA CARDIACA;					
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;					
RT "Isolation, identification and synthesis of locustamyotropin II, an					
RT additional neuropeptide of Locusta migratoria. Member of the					
RT cephalomyotropic peptide family."					
RL Insect Biochem. 20:479-484(1990).					
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY					
CC -!- (MYOTROPIC ACTIVITY).					
CC -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.					
DR INTERPRO: IPR001484; -.					
DR PROSITE: PS00539; PYROKININ; 1.					
KW Neuropeptide; Amidation; Pyrokinin.					
FT MOD_RES 8					
SQ SEQUENCE 8 AA; 934 MW; 26341771A9CAA87B CRC64;					

Query Match 36.4%; Score 20; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred No. 8.9e+04;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DGSFTQPL 9					
Db 1 EGDFTPL 8					

RESULT 2					
ANGL_BOTJA					
ID ANGL_BOTJA	STANDARD;	PRT;	10 AA.		
AC Q10581;					
DT 01-OCT-1996 (Rel. 34, Created)					
DT 01-OCT-1996 (Rel. 34, Last sequence update)					
DT 15-DEC-1998 (Rel. 37, Last annotation update)					
DE ANGIOTENSIN-LIKE PEPTIDE I (FRAGMENT).					
OS Bothrops jararaca (Jararaca).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;					
OC Viperidae; Crotalinae; Bothrops.					
RN [1]					
RP SEQUENCE.					
RC TISSUE=PLASMA;					
RX MEDLINE=96208932; Pubmed=8829801;					
RA Borgheresi R.A.M.B., Dalle Lucca J., Carmona E., Picarelli Z.P.,					

```

RT "Isolation and identification of angiotensin-like peptides from the
RT plasma of the snake Bothrops jararaca."
RL Comp. Biochem. Physiol. 113B:467-473(1996).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR INTERPRO: IPR000215;
DR PROSITE: PS00284; SERPIN; PARTIAL.
KW Vasoconstrictor; Plasma; Serpin.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1308 MW; CEF50DD761F2DB42 CRC64;

Query Match 34.5%; Score 19; DB 1; Length 10;
Best Local Similarity 33.3%; Pred. No. 3.7e+02;
Matches 3; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 DGSFTQPLX 10
Db 1 DRVYVHPFY 9

RESULT 3
LANC_CARUI STANDARD; PRT; 7 AA.
AC P36960;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE LANTIBIOTIC CARNOCIN U149 (FRAGMENT).
OS Carnobacterium sp. (strain U149).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Carnobacterium.
RN [1]
RP SEQUENCE.
RX MEDLINE=92312768; PubMed=1622206;
RA Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,
RA Nes I.F.;
RT "Purification and characterization of a new bacteriocin isolated from
RT a Carnobacterium sp.;"
RL Appl. Environ. Microbiol. 58:1417-1422(1992).
CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC).
CC ACTIVE ON GRAM-POSITIVE BACTERIA.
KW Antibiotic; Lantibiotic.
FT NON_TER 7
SQ SEQUENCE 7 AA; 786 MW; 741776D05B05B810 CRC64;

Query Match 32.7%; Score 18; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 8.9e+04;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GSFTQP 8
Db 1 GSEIQP 6

RESULT 4
ALL6_CVDPO STANDARD; PRT; 8 AA.
AC P82157;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYDIASTATIN 6
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricoidae; Tortricidae; Olethreutinae; Cydia.
RN [1]
RP SEQUENCE.
RC TISSUE=LARVA;
RX MEDLINE=98054539; PubMed=9392829;
RA Davey H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;

```

```

RT "Lepidopteran peptides of the allatostatin superfamily.;"
RL Peptides 18:1301-1309(1997).
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 936 MW; 0B2879C45B573767 CRC64;

Query Match 32.7%; Score 18; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 PLY 10
Db 2 PLY 4

RESULT 5
VP19_HSV1K STANDARD; PRT; 6 AA.
AC P23210;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE CAPSID ASSEMBLY AND DNA MATURATION PROTEIN (VIRION PROTEIN UL38)
DE (CAPSID PROTEIN VP19C) (FRAGMENT).
GN UL38
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91101287; PubMed=1846198;
RA Flanagan W.M., Papavasiliou A.G., Rice M., Hecht L.B.,
RA Silverstein S., Wagner E.K.;
RT "Analysis of the herpes simplex virus type 1 promoter controlling the
RT expression of UL38, a true late gene involved in capsid assembly.;"
RL J. Virol. 65:769-786(1991).
CC -1- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
CC EMBEDDED. BINDS DNA.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP19C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M57646; AAA45830.1;
KW Capsid assembly; Coat protein; DNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;

Query Match 29.1%; Score 16; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 8.9e+04;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 TOPL 9
Db 3 TNPL 6

RESULT 6
LCK4_LEUMA STANDARD; PRT; 8 AA.
AC P21143;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE LEUCOKININ IV (L-IV).

```



OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidea; Blaberidae; Leucophaea.  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE-HEAD;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Primary structure and synthesis of two additional neuropeptides  
 from Leucophaea maderae: members of a new family of  
 Cephalomyotropins.";  
 RL Comp. Biochem. Physiol. 84C:271-276(1986).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 ACTIVITY OF COCKROACH PROTEIN (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 906 MW; DC6365B1E9D5BDDA CRC64;

Query Match 29.1%; Score 16; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 DGSF 5  
 DB 1 DAS 4

RESULT 7  
 LPK LEUMA STANDARD; PRT; 8 AA.  
 AC P13049;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE LEUCOPYROKININ (LPK) (LEM-PK).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidea; Blaberidae; Leucophaea.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=86269041; PubMed=3015140;  
 RA Nachman R.J., Holman G.M., Cook B.J.;  
 RT "Active fragments and analogs of the insect neuropeptide  
 leucopyrokinin: structure-function studies.";  
 RL Biochem. Biophys. Res. Commun. 137:936-942(1986).  
 RN [2]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE-HEAD;  
 RX MEDLINE=87052651; PubMed=2877794;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Primary structure and synthesis of a blocked myotropic  
 neuropeptide isolated from the cockroach, Leucophaea maderae.";  
 RL Comp. Biochem. Physiol. 85C:219-224(1986).  
 CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY  
 (MYOTROPIC ACTIVITY).  
 CC -!- MISCELLANEOUS: AN ANALOG WITHOUT THE N-TERMINAL PCA RESIDUE WAS  
 SYNTHESIZED AND FOUND TO EXHIBIT GREATER ACTIVITY (144%) THAN THE  
 PARENT NEUROPEPTIDE. THE PORTION OF THE SEQUENCE OF LPK MOST  
 CRITICAL FOR THE MYOTROPIC PROPERTIES IS LIMITED TO THE  
 PENTAPEPTIDE FRAGMENT FTPLR.  
 CC -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
 CC PIR; A23967; A23967.  
 DR INTERPRO: IPR001484; .  
 DR PROSITE: PS00539; PYROKININ; 1.  
 KW Neuropeptide; Amidation; Pyrokinin.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 949 MW; 92341771A9D5A1B6 CRC64;

Query Match 29.1%; Score 16; DB 1; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 SFTQPL 9  
 DB 3 SFTPL 8

RESULT 8  
 TOP1\_SALTY STANDARD; PRT; 10 AA.  
 ID AC P40686;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE DNA TOPOISOMERASE I (EC 5.99.1.2) (OMEGA-PROTEIN) (RELAXING ENZYME)  
 DE (UNTWISTING ENZYME) (SWIVELASE) (FRAGMENT).  
 GN TOPA.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87194810; PubMed=3032952;  
 RA Ostrowski J., Jagura-Burdzy G., Kredich N.M.;  
 RT "DNA sequences of the cysB regions of Salmonella typhimurium and  
 Escherichia coli.";  
 RL J. Biol. Chem. 262:5999-6005(1987).  
 CC -!- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE  
 CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.  
 CC -!- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED  
 DNA, FOLLOWED BY PASSAGE AND REJOINING.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA  
 BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN  
 WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS  
 AT ONE END OF THE ENZYME-SEVERED DNA STRAND.  
 CC -!- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE  
 FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; M15040; AAA27044.1; .  
 DR HSSP; P06612; LYUA.  
 DR STYGENE; SG10515; TOPA.  
 DR INTERPRO: IPR000380; .  
 DR PROSITE: PS00396; TOPOISOMERASE\_I\_PROK; PARTIAL.  
 KW Isomerase; Topoisomerase; DNA-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 10 AA; 1145 MW; 2C25B67B02D37338 CRC64;

Query Match 27.3%; Score 15; DB 1; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 2.3e+03;  
 Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 DGSFTQ 7  
 DB 2 DGKWE 7

RESULT 9  
 MOSF CLYJA STANDARD; PRT; 9 AA.  
 ID AC P19853;  
 DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DE 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE [PHE-6]-MOSACT.  
 OS Clypeaster japonicus (Sand dollar).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Gnathostomata; Clypeasteroidea;  
 OC Clypeasteridae; Clypeaster.  
 RN [1]  
 RN SEQUENCE.  
 RC TISSUE=EGG JELLY;  
 RA Suzuki N., Kurita M., Yoshino K.I., Kajiuura H., Nomura K.,  
 RA Yamaguchi M.;  
 RA "Purification and structure of mosaic and its derivatives from the  
 RT egg jelly of the sea urchin Clypeaster japonicus.";  
 RL Zool. Sci. 4:649-656(1987).  
 CC -1- FUNCTION: STIMULATES SPERM RESPIRATION AND MOTILITY.  
 CC PIR: JN0027; JN0027.  
 DR PIR: JN0027; JN0027.  
 SQ SEQUENCE 9 AA; 924 MW; 93245729CDC5BAB5 CRC64;

Query Match 25.5%; Score 14; DB 1; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDGSF 5  
 Db 2 SDSAF 6

RESULT 10  
 ID UP3\_HUMAN STANDARD; PRT; 9 AA.  
 AC P30089;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 11) (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RN SEQUENCE.  
 RC TISSUE=PLASMA;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochrassner D.F.;  
 RT "Plasma protein map: an update by microsequencing.";  
 RL Electrophoresis 13:707-714(1992).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.6, ITS MW IS: 46 KDA.  
 DR SWISS-2DPAGE; P30089; HUMAN.  
 FT NON\_TER 1 1  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1056 MW; 26F2B1BAF769C737 CRC64;

Query Match 25.5%; Score 14; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 PLY 10  
 Db 2 PLF 4

RESULT 11  
 ID ANGT\_BOVIN STANDARD; PRT; 10 AA.  
 AC P01017;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-2000 (Rel. 01, Last annotation update)

DE ANGIOTENSINOGEN [CONTAINS: ANGIOTENSIN I; ANGIOTENSIN II] (FRAGMENT).  
 GN AGT.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 RN [1]  
 RN SEQUENCE.  
 RA Elliott D.F., Peart W.S.;  
 RT "The amino acid sequence in a hypertensin.";  
 RL Biochem. J. 65:246-254(1957).  
 CC -1- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN  
 CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN  
 CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE  
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT  
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL  
 CC BALANCE OF BODY FLUIDS.  
 CC -1- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.  
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
 CC PIR: A01250; A01250.  
 DR PIR: A0345; A90345.  
 DR INTERPRO: IPR000215;  
 DR PROSITE: PS00284; SERPIN; PARTIAL.  
 KW Vasoconstrictor; Plasma; Serpin.  
 FT PEPTIDE 1 10 ANGIOTENSIN I.  
 FT PEPTIDE 1 8 ANGIOTENSIN II.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1282 MW; CEEFBD761F2DB42 CRC64;

Query Match 25.5%; Score 14; DB 1; Length 10;  
 Best Local Similarity 22.2%; Pred. No. 3.7e+03;  
 Matches 2; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 DGSFTQPLY 10  
 Db 1 DRVYVHPFH 9

RESULT 12  
 ID TUFT\_HUMAN STANDARD; PRT; 4 AA.  
 AC P01858;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE PHAGOCYTOSIS-STIMULATING PEPTIDE (TUFTSIN).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RN SEQUENCE.  
 RX MEDLINE=72187087; PubMed=4112769;  
 RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;  
 RT "The characteristics, isolation and synthesis of the phagocytosis  
 RT stimulating peptide tuftsin.";  
 RL Biochem. Biophys. Res. Commun. 47:172-179(1972).  
 RN [2]  
 RN IMMUNOGLOBULIN CLASS.  
 RX MEDLINE=68091045; PubMed=4169272;  
 RA Fidalgo B.V., Najjar V.A.;  
 RT "The physiological role of the lymphoid system. VI. The stimulatory  
 RT effect of leucophilic gamma globulin (leucokinin) on the phagocytic  
 RT activity of human polymorphonuclear leucocyte.";  
 RL Biochemistry 6:3386-3392(1967).  
 CC -1- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSIBLY TO THE  
 CC CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE  
 CC MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN.  
 CC TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC  
 CC ACTIVITY OF NEUTROPHILS.  
 DR PIR: A02147; A02147.  
 DR MIM: 191150;  
 SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;

Query Match 23.6%; Score 13; DB 1; Length 4;  
Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TOP 8  
I: I  
Db 1 TRP 3

## RESULT 13

FARP\_MONEX STANDARD; PRT; 6 AA.  
AC P41966;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE FMRFAMIDE-LIKE NEUROPEPTIDE GNFRF-AMIDE.  
OS Moniezia expansa (Sheep tapeworm).  
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Cestoda; Eucestoda;  
OC Cyclophyllidae; Anoplocephalidae; Moniezia.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93312289; PubMed=8323531;  
RA Maule A.G., Shaw C., Halton D.W., Thim L.;  
RT "GNFRFamide: a novel FMRFamide-immunoreactive peptide isolated from  
the sheep tapeworm, Moniezia expansa.";  
RL Biochem. Biophys. Res. Commun. 193:1054-1060(1993).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 6 6 AMIDATION.  
SQ SEQUENCE 6 AA; 787 MW; 69D409C9C4481000 CRC64;

Query Match 23.6%; Score 13; DB 1; Length 6;  
Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GSF 5  
I: I  
Db 1 GNF 3

## RESULT 14

MNP1\_LEPDE STANDARD; PRT; 7 AA.  
AC P42984;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE MYOTROPIC NEUROPEPTIDE 1 (LED-MNP-1).  
OS Leptinotarsa decemlineata (Colorado potato beetle).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
OC Cucujiformia; Chrysomelidae; Chrysomelinae; Leptinotarsa.  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE-HEAD;  
RX MEDLINE=95380343; PubMed=7651886;  
RA Spittaels K., Vankeerberghen A., Schoofs L., Torrekens S.,  
Grauwels L., van Leuven F., de Loof A.;  
RT "Identification, characterization, and immunological localization of  
a novel myotropic neuropeptide in the Colorado potato beetle,  
Leptinotarsa decemlineata.";  
RL Peptides 16:365-374(1995).  
CC -!- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE  
OVIDUCT.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7 AMIDATION.  
SQ SEQUENCE 7 AA; 705 MW; 6DD73768745B5DB0 CRC64;

Query Match 23.6%; Score 13; DB 1; Length 7;  
Best Local Similarity 33.3%; Pred. No. 8.9e+04;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 SFTQPL 9  
I: I  
Db 1 AYNGL 6

## RESULT 15

GLUR\_HUMAN STANDARD; PRT; 8 AA.  
AC P02729;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE URINE GLYCOPEPTIDE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=72062338; PubMed=5126885;  
RA Lote C.J., Weiss J.B.;  
RT "Identification in urine of a low-molecular-weight highly polar  
glycopeptide containing cysteinyl-galactose.";  
RL Biochem. J. 123:25P-25P(1971).  
CC -!- FUNCTION: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS PEPTIDE  
IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN  
ATTRIBUTED. AN ERYTHROCYTE MEMBRANE GLYCOPEPTIDE HAVING A  
SIMILAR STRUCTURE HAS ALSO BEEN FOUND.  
CC PIR; A03188; XGHUEU.  
DR Glycoprotein. 1  
FT CARBOHYD 1 S-LINKED (GAL. .).  
SQ SEQUENCE 8 AA; 855 MW; C2D87AA1F561EBIE CRC64;

Query Match 23.6%; Score 13; DB 1; Length 8;  
Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DGS 4  
I: I  
Db 6 DGA 8

Search completed: January 17, 2001, 13:41:34  
Job time: 103 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:39:31 ; Search time 19.29 Seconds  
(without alignments)  
60.761 Million cell updates/sec

Title: US-08-765-837-3  
Perfect score: 55  
Sequence: 1 TDGSFTQPLY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 815

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	34.5	9	8 Q9TLM7	Q9TLM7 laurencia v
2	19	34.5	9	11 Q9QWT0	Q9QWT0 mus musculus
3	19	34.5	10	10 Q9S905	Q9S905 glycine max
4	18	32.7	9	2 Q33914	Q53914 streptomyces
5	17	30.9	9	11 Q9QVH9	Q9QVH9 mus sp. sup
6	16	29.1	8	2 Q9RQ57	Q9RQ57 buchnera ap
7	16	29.1	8	2 Q9RQ49	Q9RQ49 buchnera ap
8	16	29.1	8	7 Q29810	Q29810 homo sapien
9	16	29.1	9	10 Q9S8J8	Q9S8J8 oryza sativ
10	15	27.3	5	13 P82073	P82073 littoria rub
11	15	27.3	5	13 P82099	P82099 littoria rub
12	15	27.3	7	2 O50556	O50556 actinobacil
13	15	27.3	10	4 Q9UCQ8	Q9UCQ8 homo sapien
14	14	25.5	7	5 Q9VYN9	Q9VYN9 drosophila
15	14	25.5	8	6 Q9T778	Q9T778 canis famli
16	14	25.5	8	10 Q40530	Q40530 nicotiana t
17	14	25.5	9	12 Q9IBY9	Q9IBY9 kaposi's sa
18	14	25.5	10	4 O00493	O00493 homo sapien
19	14	25.5	10	11 Q9R0Z0	Q9R0Z0 mus musculus

20	14	25.5	10	12	Q86324	Q86324 rous sarcom
21	14	25.5	10	12	Q86325	Q86325 rous sarcom
22	14	25.5	10	12	Q86326	Q86326 rous sarcom
23	14	25.5	10	12	Q9J168	Q9J168 tt virus. o
24	13	23.6	7	10	P93233	P93233 lycopersico
25	13	23.6	8	3	P87225	P87225 saccharomyc
26	13	23.6	9	4	Q14277	Q14277 homo sapien
27	13	23.6	9	6	Q9TRSO	Q9TRSO oryctolagus
28	13	23.6	10	2	Q9R5N1	Q9R5N1 clostridium
29	13	23.6	10	11	Q9QVF7	Q9QVF7 rattus sp.
30	13	23.6	10	12	Q86580	Q86580 simian para
31	13	23.6	10	12	Q9QX97	Q9QX97 polyomaviru
32	13	23.6	10	12	Q9QX77	Q9QX77 polyomaviru
33	13	23.6	10	12	Q9QX57	Q9QX57 polyomaviru
34	13	23.6	10	12	Q9QX37	Q9QX37 polyomaviru
35	13	23.6	10	12	Q9QX17	Q9QX17 polyomaviru
36	13	23.6	10	12	Q9QW97	Q9QW97 polyomaviru
37	13	23.6	10	12	Q9QW77	Q9QW77 polyomaviru
38	13	23.6	10	12	Q9QW57	Q9QW57 polyomaviru
39	13	23.6	10	12	Q9QW37	Q9QW37 polyomaviru
40	13	23.6	10	12	Q9QW17	Q9QW17 polyomaviru
41	13	23.6	10	12	Q9QV97	Q9QV97 polyomaviru
42	13	23.6	10	12	Q9QV77	Q9QV77 polyomaviru
43	12	21.8	8	2	Q9S6D5	Q9S6D5 escherichia
44	12	21.8	8	2	Q9R9E0	Q9R9E0 bacillus su
45	12	21.8	8	4	Q15888	Q15888 homo sapien

#### ALIGNMENTS

RESULT 1  
Q9TLM7  
ID Q9TLM7 PRELIMINARY: PRT: 9 AA.  
AC Q9TLM7  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE SMALL SUBUNIT  
DE (FRAGMENT).  
GN RBCL.  
OS Laurencia viridis.  
OG Chloroplast.  
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Rhodomelaceae;  
OC Laurencia.  
OX NCBI\_taxid=99904;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lewis S.M., Valdes F., Frias I., Hernandez-Gonzalez M.C.,  
RA Gil-Rodriguez M.C., Gacesa P.,  
RT "Analysis of rDNA and RUBISCO spacer sequences of the genera Laurencia  
and Osmundea in the Canary Islands.",  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF081272; AAD53999.1; -  
KW Chloroplast.  
FT NON\_TER  
SQ SEQUENCE 9 AA; 1040 MW; 7080DIA866C1A734 CRC64;

Query Match 34.5%; Score 19; DB 8; Length 9;  
Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDGSFT 6  
| : | :  
Db 4 TQGTFS 9

RESULT 2  
Q9QWT0  
ID Q9QWT0 PRELIMINARY: PRT: 9 AA.  
AC Q9QWT0  
DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE PROTEINASE 3 (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SV129 D3;  
 RA Lutz P.G., Moog-Lutz C., Houzel-Charavel A., Cayre Y.E.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ007030; CAA07429.1;  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 937 MW; C91E75A77B45B87D CRC64;

Query Match 34.5%; Score 19; DB 11; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GSFTOP 8  
 ||: |  
 Db 3 GSPSP 8

RESULT 3  
 O9S905 PRELIMINARY; PRT; 10 AA.  
 AC Q9S905;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE S3 PEPTIDE.  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE.  
 RA Hirano H., Kagawa H., Okubo K.;  
 RL Phytochemistry 31:731-735(1992).  
 SQ SEQUENCE 10 AA; 975 MW; 3C733271A879D1B7 CRC64;

Query Match 34.5%; Score 19; DB 10; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DGSFTOPL 9  
 |.: |  
 Db 2 DPTFGTPL 9

RESULT 4  
 Q53914 PRELIMINARY; PRT; 9 AA.  
 AC Q53914;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE CURD PROTEIN (FRAGMENT).  
 OS Streptomyces cyaneus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=1904;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bergh S.T., Uhlen M.;  
 RL Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M33704; AAA26724.1;  
 FT NON\_TER 1

SQ SEQUENCE 9 AA; 1027 MW; 995BDDDDCA140AB1 CRC64;  
 Query Match 32.7%; Score 18; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TDGSFTQP 8  
 ||: |  
 Db 2 TDRRAQP 9

RESULT 5  
 Q9QVH9 PRELIMINARY; PRT; 9 AA.  
 AC Q9QVH9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE SUPERINDUCIBLE PROTEIN 24, SIP24=CYCLOPHILIN HOMOLOG, PEAK C.  
 OS Mus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10095;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92031730; PubMed=1932134;  
 RA Davis T.R., Tabatabai L., Bruns K., Hamilton R.T., Nilsen-Hamilton M.;  
 RT "Basic fibroblast growth factor induces 3T3 fibroblasts to synthesize  
 and secrete a cyclophilin-like protein and beta 2-microglobulin.";  
 RL Biochim. Biophys. Acta 1095:145-152(1991).  
 SQ SEQUENCE 9 AA; 963 MW; D6D27731A9D6D87D CRC64;

Query Match 30.9%; Score 17; DB 11; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GSFT 6  
 |.: |  
 Db 3 GQFT 6

RESULT 6  
 Q9RQ57 PRELIMINARY; PRT; 8 AA.  
 ID Q9RQ57;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE NIFS PROTEIN HOMOLOG (FRAGMENT).  
 GN NIFS.  
 OS Buchnera aphidicola.  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=9;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20022990; PubMed=1055290;  
 RA Clark M.A., Moran N.A., Baumann P.;  
 RT "Sequence evolution in bacterial endosymbionts having extreme base  
 compositions.";  
 RL Mol. Biol. Evol. 16:1586-1598(1999).  
 DR EMBL; AF130812; AAF13797.1;  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 980 MW; F3A73B504771A336 CRC64;

Query Match 29.1%; Score 16; DB 2; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 PLY 10  
 |.: |

```

Db      4 PIY 6

RESULT  7
Q9RQ49 ID Q9RQ49 PRELIMINARY; PRT; 8 AA.
AC Q9RQ49;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE NIFS PROTEIN HOMOLOG (FRAGMENT).
GN NIFS.
OS Buchnera aphidicola.
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=9;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20022990; PubMed=10555290;
RA Clark M.A., Moran N.A., Baumann P.;
RT "Sequence evolution in bacterial endosymbionts having extreme base
RT compositions.";
RL MOL. Biol. Evol. 16:1586-1598(1999).
DR EMBL; AF130814; AAF13805.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 992 MW; F3A73B5047672336 CRC64;

Query Match 29.1%; Score 16; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 PIY 10
Db 4 PIY 6

RESULT  8
Q29810 ID Q29810 PRELIMINARY; PRT; 8 AA.
AC Q29810;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HLA CLASS II DR-BETA CHAIN (FRAGMENT).
GN HLA-DRB5*15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96175156; PubMed=9110934;
RA Svensson A.C., Setterblad N., Pihlgren U., Rask L., Andersson G.;
RT "Evolutionary relationship between human major histocompatibility
RT complex HLA-DR haplotypes.";
RL Immunogenetics 43:304-314(1996).
DR EMBL; X88792; CAA61271.1; -.
KW MHC.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 909 MW; 6DC6C1A9CAA731A8 CRC64;

Query Match 29.1%; Score 16; DB 7; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 FTQ 7
Db 5 FTQ 7

RESULT  9
Q9S8J8 ID Q9S8J8 PRELIMINARY; PRT; 9 AA.
AC Q9S8J8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE ORYZATENSIN-BIOACTIVE PEPTIDE.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE.
RX MEDLINE=95102521; PubMed=7804141;
RA Takahashi M., Moriguchi S., Yoshikawa M., Sasaki R.;
RT "Isolation and characterization of Oryzatenin: a novel bioactive
RT peptide with ileum-contracting and immunomodulating activities derived
RT from rice albumin.";
RL Biochem. Mol. Biol. Int. 33:1151-1158(1994).
SQ SEQUENCE 9 AA; 1093 MW; 0E8C67377B56877B CRC64;

Query Match 29.1%; Score 16; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 PIY 10
Db 3 PIY 5

RESULT 10
P82073 ID P82073 PRELIMINARY; PRT; 5 AA.
AC P82073;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE RUBELLIDIN 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RX TISSUE=SKIN SECRETION;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:0-0(1999).
CC -!- FUNCTION: CARRIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;

Query Match 27.3%; Score 15; DB 13; Length 5;
Best Local Similarity 75.0%; Pred. No. 3.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GSFT 6
Db 2 GSFT 5

RESULT 11
P82099 ID P82099 PRELIMINARY; PRT; 5 AA.
AC P82099;
DT 01-MAY-2000 (TREMBlrel. 13, Created)

```

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE ELECTRIN 3.  
 OS Litoria rubella (Desert tree frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 OC Litoria.  
 OX NCBI\_TaxID=104895;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=SKIN SECRETION;  
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
 RT "Peptides from the skin glands of the Australian buzzing tree frog  
 RT Litoria electrica. Comparison with the skin peptides from Litoria  
 RT rubella.";  
 RL Aust. J. Chem. 52:0-0(1999).  
 KW Amphibian skin; Amidation.  
 FT MOD\_RES 5 5  
 SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;  
 AMIDATION.

Query Match 27.3%; Score 15; DB 13; Length 5;  
 Best Local Similarity 40.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 FTQPL 9  
 Db 1 FVHPM 5

RESULT 12  
 O50556 PRELIMINARY; PRT; 7 AA.  
 AC O50556;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE GLYA (FRAGMENT).  
 GN GLYA.  
 OS Actinobacillus actinomycetemcomitans  
 OS (Haemophilus actinomycetemcomitans).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Actinobacillus.  
 OX NCBI\_TaxID=714;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33384;  
 RX MEDLINE=96355846; PubMed=8751884;  
 RA Kolodrubetz D., Spltznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,  
 RA Kraig E.;  
 RT "cis Elements and trans factors are both important in strain-specific  
 RT regulation of the leukotoxin gene in Actinobacillus  
 RT actinomycetemcomitans.";  
 RL Infect. Immun. 64:3451-3460(1996).  
 DR EMBL; U51862; AAB88721.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 27.3%; Score 15; DB 2; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 PLY 10  
 Db 4 PLY 6

RESULT 13  
 O9UC08 PRELIMINARY; PRT; 10 AA.  
 ID O9UC08;  
 AC O9UC08;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE AUTOTAXIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC MEDLINE=92129337; PubMed=1733949;  
 RA Stracke M.L., Kruttsch H.C., Unsworth E.J., Arestad A., Cioce V.,  
 RA Schiffmann E., Liotta L.A.;  
 RT "Identification, purification, and partial sequence analysis of  
 RT autotaxin, a novel motility-stimulating protein.";  
 RL J. Biol. Chem. 267:2524-2529(1992).  
 SQ SEQUENCE 10 AA; 1171 MW; 736F44577AF1B2CB CRC64;

Query Match 27.3%; Score 15; DB 4; Length 10;  
 Best Local Similarity 60.0%; Pred. No. 1.1e+04;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 TQPLY 10  
 Db 5 TQPNY 9

RESULT 14  
 Q9VYN9 PRELIMINARY; PRT; 7 AA.  
 ID Q9VYN9;  
 AC Q9VYN9;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE CG18256 PROTEIN (FRAGMENT).  
 GN CG18256.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=107311132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,



RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turnek R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003488; AAF48153.1; -.  
 DR FLYBASE; FBgn0030380; CG18256.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 817 MW; 69C737644731A740 CRC64;

Query Match 25.5%; Score 14; DB 5; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 8 PLX 10  
 Db 5 PLF 7

RESULT 15  
 QSTT78  
 ID Q9TT78 PRELIMINARY; PRT; 8 AA.  
 AC Q9TT78;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
 DE THYMIDYLATE SYNTHASE (FRAGMENT).  
 GN TS.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Brouillette J.A., Andrew J.R., Venta P.J.;  
 RT "Estimate of Nucleotide Diversity in Dogs Using a Pool-and-Sequence  
 Method to Identify Single Nucleotide Polymorphisms."  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF202073; AAF20918.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 899 MW; 6731A1E059CAA867 CRC64;

Query Match 25.5%; Score 14; DB 6; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 GSFTQPL 9  
 Db 2 GDFIHTL 8

Search completed: January 17, 2001, 13:41:22  
 Job time: 111 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:38:15 ; Search time 12.2 seconds  
(without alignments)  
14.719 Million cell updates/sec

Title: US-08-765-837-3  
Perfect score: 55  
Sequence: 1 TDGFTQPLY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 52832

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2.6/ptodata/2/1aa/5\_COMB.pep:\*  
4: /cgn2.6/ptodata/2/1aa/PTUS\_COMB.pep:\*  
5: /cgn2.6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	27	49.1	8	2	US-08-985-090-14
2	27	49.1	8	3	US-08-159-339A-1226
3	27	49.1	8	3	US-09-165-543-34
4	27	49.1	10	3	US-08-159-339A-109
5	22	40.0	6	2	US-08-482-228-187
6	22	40.0	6	3	US-08-482-528-187
7	22	40.0	10	1	US-08-190-788A-64
8	22	40.0	10	1	US-08-383-474B-69
9	22	40.0	10	1	US-08-465-391A-64
10	22	40.0	10	2	US-08-464-538B-64
11	22	40.0	10	2	US-08-463-076E-108
12	22	40.0	10	4	PCT-US95-11127-9
13	21	38.2	6	2	US-08-482-228-8
14	21	38.2	6	3	US-08-482-528-8
15	21	38.2	7	3	US-09-173-941-100
16	21	38.2	7	5	5208144-38
17	21	38.2	7	5	5208144-39
18	21	38.2	8	1	US-08-204-656B-14
19	21	38.2	8	1	US-08-470-702-14
20	21	38.2	8	1	US-08-467-831-14
21	21	38.2	8	4	PCT-US93-01669-1
22	21	38.2	8	4	PCT-US94-01321-46
23	21	38.2	9	1	US-08-486-057B-15
24	21	38.2	9	2	US-08-789-588-15
25	21	38.2	10	1	US-08-486-057B-17
26	21	38.2	10	2	US-08-541-033A-6
27	21	38.2	10	2	US-08-789-588-17
28	21	38.2	10	2	US-08-602-725-8

29 21 38.2 10 2 US-08-828-451-6  
30 20 36.4 6 2 US-08-482-228-185  
31 20 36.4 6 3 US-08-482-528-185  
32 20 36.4 6 4 PCT-US93-01669-29  
33 20 36.4 7 2 US-08-482-228-28  
34 20 36.4 7 3 US-08-482-528-28  
35 20 36.4 8 3 US-08-863-813A-20  
36 20 36.4 8 5 5514646-17  
37 20 36.4 9 1 US-08-297-731-7  
38 20 36.4 9 1 US-08-426-819A-20  
39 20 36.4 9 4 PCT-US95-10793-7  
40 20 36.4 10 1 US-08-297-731-8  
41 20 36.4 10 4 PCT-US95-10793-8  
42 20 36.4 10 5 5210075-47  
43 19 34.5 6 1 US-07-865-216B-1  
44 19 34.5 6 2 US-08-482-228-136  
45 19 34.5 6 2 US-08-482-228-138

## ALIGNMENTS

RESULT 1  
US-08-985-090-14  
; Sequence 14, Application US/08985090  
; Patent No. 5855893  
; GENERAL INFORMATION:  
; APPLICANT: Andrew D.J. Goodearl  
; TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/985,090  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jean M. Silveri  
; REGISTRATION NUMBER: 39,030  
; REFERENCE/DOCKET NUMBER: MNI-032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-985-090-14

Query Match 49.1%; Score 27; DB 2; Length 8;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GSFTQPLY 10  
|:| |  
DB 1 GAFCIPLY 8

## RESULT 2

US-08-159-339A-1226  
; Sequence 1226, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Celis, Esteban  
; TITLE OF INVENTION: HLA Binding peptides and Their  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993  
; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-005030US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1226:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-159-339A-1226

Query Match 49.1%; Score 27; DB 3; Length 8;  
Best Local Similarity 71.4%; Pred. No. 1.3e+05;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 SFTOPLY 10  
|||  
Db 2 SFPHPY 8

## RESULT 3

US-09-165-543-34  
; Sequence 34, Application US/09165543  
; Patent No. 6093545  
; GENERAL INFORMATION:  
; APPLICANT: Andrew D.J. Goodearl and Sandra Glucksmann  
; TITLE OF INVENTION: Muscarinic Receptors and Uses Therefor  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street

; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/165,543  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/042,780  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elizabeth A. Hanley  
; REGISTRATION NUMBER: 33,505  
; REFERENCE/DOCKET NUMBER: MNI-032CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-165-543-34

Query Match 49.1%; Score 27; DB 3; Length 8;  
Best Local Similarity 62.5%; Pred. No. 1.3e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GSFTQPLY 10  
|||  
Db 1 GAFCIPLY 8

## RESULT 4

US-08-159-339A-109  
; Sequence 109, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Celis, Esteban  
; TITLE OF INVENTION: HLA Binding peptides and Their  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746

FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-0050300S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:

INFORMATION FOR SEQ ID NO: 109:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-109

Query Match 49.1%; Score 27; DB 3; Length 10;  
Best Local Similarity 71.4%; Pred. No. 30;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 SFTQPLY 10

Db 4 SFFHPLY 10

RESULT 5

US-08-482-228-187  
Sequence 187, Application US/08482228  
Patent No. 5968753

GENERAL INFORMATION:

APPLICANT: Tseng-Law, Janet  
APPLICANT: Kobori, Joan A.  
APPLICANT: Al-Abdaly, Fahad A.  
APPLICANT: Guillermo, Roy  
APPLICANT: Helgerson, Sam L.

TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL  
NUMBER OF SEQUENCES: 215  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Janice Guthrie, Ph.D.  
STREET: P.O. Box 15210  
CITY: Irvine  
STATE: California  
COUNTRY: USA  
ZIP: 92713-5210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,228  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: Guthrie, Janice

REGISTRATION NUMBER: 35,170  
REFERENCE/DOCKET NUMBER: IT-4630CIP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (714) 440-5353

TELEFAX: (714) 553-1952

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-482-228-187

Query Match 40.0%; Score 22; DB 2; Length 6;  
Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 FTQPL 9

Db 2 YTOPI 6

RESULT 6

US-08-482-528-187

Sequence 187, Application US/08482528  
Patent No. 6017719

GENERAL INFORMATION:

APPLICANT: Tseng-Law, Janet

APPLICANT: Kobori, Joan A.

APPLICANT: Al-Abdaly, Fahad A.

APPLICANT: Guillermo, Roy

APPLICANT: Helgerson, Sam L.

APPLICANT: Deans, Robert J.

TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL

NUMBER OF SEQUENCES: 215

CORRESPONDENCE ADDRESS:

ADDRESSEE: Janice Guthrie, Ph.D.

STREET: P.O. Box 15210

CITY: Irvine

STATE: California

COUNTRY: USA

ZIP: 92713-5210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,528

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Guthrie, Janice

REGISTRATION NUMBER: 35,170

REFERENCE/DOCKET NUMBER: IT-4630CIP4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (714) 440-5353

TELEFAX: (714) 553-1952

INFORMATION FOR SEQ ID NO: 187:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-482-528-187

Query Match 40.0%; Score 22; DB 3; Length 6;  
Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 FTQPL 9

Db 2 YTOPI 6

RESULT 7

US-08-190-788A-64

Sequence 64, Application US/08190788A

Patent No. 5608035

GENERAL INFORMATION:

APPLICANT: Yanofsky, Stephen D.

APPLICANT: Barrett, Ronald W.  
APPLICANT: Baldwin, David N.  
APPLICANT: Jacobs, Jeff W.  
TITLE OF INVENTION: Peptides and Compounds That Bind to the  
TITLE OF INVENTION: IL-1 Receptor  
NUMBER OF SEQUENCES: 312  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies N.V.  
STREET: 4001 Miranda Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,788A  
FILING DATE: 02-FEB-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,567  
FILING DATE: 05-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 1019.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-190-788A-64

Query Match 40.0%; Score 22; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GSFTQP 8  
||: ||  
Db 1 GSWEQP 6

RESULT 8  
US-08-383-474B-69  
Sequence 69, Application US/08383474B  
Patent No. 5767234  
GENERAL INFORMATION:  
APPLICANT: Yanofsky, Stephen D.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Baldwin, David N.  
TITLE OF INVENTION: Peptides and Compounds That Bind to  
TITLE OF INVENTION: the IL-1 Receptor  
NUMBER OF SEQUENCES: 314  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend & Townsend & Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/383,474B  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,788  
FILING DATE: 02-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 1019.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-383-474B-69  
Query Match 40.0%; Score 22; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GSFTQP 8  
||: ||  
Db 1 GSWEQP 6

RESULT 9  
US-08-465-391A-64  
Sequence 64, Application US/08465391A  
Patent No. 5786331  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Yanofsky, Stephen D.  
APPLICANT: Baldwin, David  
APPLICANT: Jacobs, Jeff W.  
APPLICANT: Bovy, Philippe R.  
APPLICANT: Leahy, Ellen M.  
APPLICANT: Pottorf, Richard S.  
TITLE OF INVENTION: Peptides and Compounds That Bind to the  
TITLE OF INVENTION: IL-1 Receptor  
NUMBER OF SEQUENCES: 405  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,391A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,474  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,788

; FILING DATE: 02-FEB-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5786331v1el, Vern  
; REGISTRATION NUMBER: 32,483  
; REFERENCE/DOCKET NUMBER: 16528A-001840/1019.2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-465-391A-64

Query Match 40.0%; Score 22; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 GSFTQP 8  
||: ||  
Db 1 GSWEQP 6

RESULT 10  
US-08-464-538B-64  
; Sequence 64, Application US/08464538B  
; Patent No. 5861476  
; GENERAL INFORMATION:  
; APPLICANT: Barrett, Ronald W.  
; APPLICANT: Yanofsky, Stephen D.  
; APPLICANT: Baldwin, David  
; APPLICANT: Jacobs, Jeff W.  
; APPLICANT: Bovy, Phillippe R.  
; APPLICANT: Leahy, Ellen M.  
; APPLICANT: Pottorf, Richard S.  
; TITLE OF INVENTION: Peptides and Compounds That Bind to the  
; TITLE OF INVENTION: IL-1 Receptor  
; NUMBER OF SEQUENCES: 402  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,538B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,474  
; FILING DATE: 01-FEB-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/190,788  
; FILING DATE: 02-FEB-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 16528A-001810  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400

; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-464-538B-64  
; Query Match 40.0%; Score 22; DB 2; Length 10;  
; Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GSFTQP 8  
||: ||  
Db 1 GSWEQP 6

RESULT 11  
US-08-463-076E-108  
; Sequence 108, Application US/08463076E  
; Patent No. 5880096  
; GENERAL INFORMATION:  
; APPLICANT: Barrett, Ronald W.  
; APPLICANT: Yanofsky, Stephen D.  
; TITLE OF INVENTION: Peptides and Compounds That Bind to the  
; TITLE OF INVENTION: IL-1 Receptor  
; NUMBER OF SEQUENCES: 392  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,076E  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Snyder, Joseph R.  
; REGISTRATION NUMBER: 39,381  
; REFERENCE/DOCKET NUMBER: 16528A-001850US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 108:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-463-076E-108

Query Match 40.0%; Score 22; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GSFTQP 8  
||: ||  
Db 1 GSWEQP 6

RESULT 12

PCT-US95-11127-9  
; Sequence 9, Application PC/TUS9511127  
; GENERAL INFORMATION:  
; APPLICANT: ROBERT WEBBER  
; TITLE OF INVENTION: DOPAMINE RECEPTOR PEPTIDES AND  
; ANTI-BODIES  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIELEN, PETERSON & LAMPE  
; STREET: 1990 N. CALIFORNIA BOULEVARD, SUITE 720  
; CITY: WALNUT CREEK  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94596  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE 5.25 INCH, 1.2 MB FOR FORMATTED  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/11127  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: NONE  
; FILING DATE: NONE  
; ATTORNEY/AGENT INFORMATION:  
; NAME: THEODORE J. BIELEN, JR.  
; REGISTRATION NUMBER: 27,420  
; REFERENCE/DOCKET NUMBER: 12068  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 937-1515  
; TELEFAX: (510) 937-1529  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PEPTIDE  
; FEATURE:  
; NAME/KEY: D2 RECEPTOR (AC-25-34)  
; LOCATION:  
; IDENTIFICATION METHOD: AMINO ACID ANALYSIS  
; OTHER INFORMATION: DOPAMINE D2 RECEPTOR PEPTIDE  
PCT-US95-11127-9  
  
Query Match 40.0%; Score 22; DB 4; Length 10;  
Best Local Similarity 40.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 TDGSFTOPLY 10  
Db 1 SDGKADRPY 10  
  
RESULT 13  
US-08-482-228-8  
; Sequence 8, Application US/08482228  
; Patent No. 5968753  
; GENERAL INFORMATION:  
; APPLICANT: Tseng-Law, Janet  
; APPLICANT: Kobori, Joan A.  
; APPLICANT: Al-Abdaly, Fahad A.  
; APPLICANT: Guillermo, Roy  
; APPLICANT: Helgerson, Sam L.  
; APPLICANT: Deans, Robert J.  
; TITLE OF INVENTION: POSITIVE AND NEGATIVE/NEGATIVE CELL  
; SELECTION MEDIATED BY PEPTIDE RELEASE  
; NUMBER OF SEQUENCES: 215  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Janice Guthrie, Ph.D.  
; STREET: P.O. Box 15210  
; CITY: Irvine  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92713-5210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,528  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Guthrie, Janice  
; REGISTRATION NUMBER: 35,170  
; REFERENCE/DOCKET NUMBER: IT-4630CIP4  
; TELECOMMUNICATION INFORMATION:

CITY: Irvine  
STATE: California  
COUNTRY: USA  
ZIP: 92713-5210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,228  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Guthrie, Janice  
REGISTRATION NUMBER: 35,170  
REFERENCE/DOCKET NUMBER: IT-4630CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (714) 440-5353  
TELEFAX: (714) 553-1952  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-228-8  
  
Query Match 38.2%; Score 21; DB 2; Length 6;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 TDGSF 5  
Db 1 TQGSF 5  
  
RESULT 14  
US-08-482-528-8  
; Sequence 8, Application US/08482528  
; Patent No. 6017719  
; GENERAL INFORMATION:  
; APPLICANT: Tseng-Law, Janet  
; APPLICANT: Kobori, Joan A.  
; APPLICANT: Al-Abdaly, Fahad A.  
; APPLICANT: Guillermo, Roy  
; APPLICANT: Helgerson, Sam L.  
; APPLICANT: Deans, Robert J.  
; TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL  
; SELECTION MEDIATED BY PEPTIDE RELEASE  
; NUMBER OF SEQUENCES: 215  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Janice Guthrie, Ph.D.  
; STREET: P.O. Box 15210  
; CITY: Irvine  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92713-5210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,528  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Guthrie, Janice  
; REGISTRATION NUMBER: 35,170  
; REFERENCE/DOCKET NUMBER: IT-4630CIP4  
; TELECOMMUNICATION INFORMATION:



; TELEPHONE: (714) 440-5353  
; TELEFAX: (714) 553-1952  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-482-528-8

Query Match 38.2%; Score 21; DB 3; Length 6;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDGSF 5  
| | | |  
DB 1 TQGSF 5

RESULT 15  
US-09-173-941-100  
; Sequence 100, Application US/09173941  
; Patent No. 6140081  
; GENERAL INFORMATION:  
; APPLICANT: BARBAS, Carlos F.  
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN  
; FILE REFERENCE: NOV0081S  
; CURRENT APPLICATION NUMBER: US/09/173,941  
; CURRENT FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 100  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: nucleotide  
; OTHER INFORMATION: codon binding sequence  
US-09-173-941-100

Query Match 38.2%; Score 21; DB 3; Length 7;  
Best Local Similarity 57.1%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TDGSFTQ 7  
| | | |  
DB 1 TSGSLTR 7

Search completed: January 17, 2001, 13:40:02  
Job time: 107 sec



orig hit - then

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:40:06 ; Search time 43.61 Seconds  
(without alignments)  
7.057 Million cell updates/sec

Title: US-08-765-837-4  
Perfect score: 48  
Sequence: 1 NQASRPYSF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 61695

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_36.\*  
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	9	R90948	Factor VIII antige
2	33	68.8	8	W32492	Helicostatin 3, wh
3	31	64.6	7	10 P90351	Peptide fragment o
4	31	64.6	7	10 P90354	Peptide fragment o
5	30	62.5	8	18 W32493	Helicostatin 4, wh
6	29	60.4	7	15 R62417	Dipteran neuropepe
7	29	60.4	7	18 W32499	Callatostatin 6, w
8	27	56.2	8	18 W32506	Pteridostatin 2, w
9	27	56.2	9	20 Y47047	Immunogenic peptid
10	25	52.1	7	10 P90859	New bradykinin ana
11	25	52.1	9	21 Y55007	Mouse histocompati
12	24	50.0	9	15 R71753	Neurotensin recept

13	23	47.9	8	20	Y48827	Membrane dipeptida
14	22	45.8	9	15	R71755	Neurotensin recept
15	22	45.8	9	20	Y55541	HLA binding plu-1
16	22	45.8	9	21	Y44591	Mouse anti-IL-18 a
17	21	43.8	5	21	W90991	N. denitrificans a
18	21	43.8	6	20	Y01578	Antigenic peptide
19	21	43.8	7	20	Y16921	Heat shock protein
20	21	43.8	8	9	P81974	New substrate 3 fo
21	21	43.8	8	15	R48278	Cytoskeleton local
22	21	43.8	8	20	W96376	Localisation signa
23	21	43.8	8	21	Y54429	Peptide which mini
24	21	43.8	9	18	W16655	Anti-cancer specif
25	21	43.8	9	20	Y55540	HLA binding plu-1
26	20	41.7	6	15	R49938	Neuropeptide repre
27	20	41.7	6	16	R83634	C.vomitorea allato
28	20	41.7	6	20	Y01575	Antigenic peptide
29	20	41.7	6	20	Y01576	Antigenic peptide
30	20	41.7	6	20	Y01565	Antigenic peptide
31	20	41.7	6	21	W80301	Synthetic citrate
32	20	41.7	7	20	W93248	Human cytochrome p
33	20	41.7	7	21	Y80302	Modified citrate s
34	20	41.7	8	15	R62418	Dipteran neuropepe
35	20	41.7	8	16	R83632	C.vomitorea allato
36	20	41.7	8	17	R98051	Pseudosterialin G,
37	20	41.7	8	18	W32507	Pteridostatin 3, w
38	20	41.7	8	21	Y77679	Hyaluronic acid (H
39	20	41.7	9	10	P90680	New antihypertensi
40	20	41.7	9	16	W21420	Alzheimer amyloid
41	20	41.7	9	17	W49476	Human leucocyte an
42	20	41.7	9	18	W30959	Timothy grass poll
43	20	41.7	9	18	W30993	Timothy grass poll
44	20	41.7	9	19	W83248	NPF motif EH domai
45	20	41.7	9	20	Y53343	Bcr-Abl epitope (a

ALIGNMENTS

RESULT 1  
R90948 R90948 standard; peptide; 9 AA.  
XX  
AC R90948;  
XX  
DT 09-SEP-1996 (first entry)  
XX  
DE Factor VIII antigenic peptide corresp. to residues Aspl777-Phel785.  
XX  
KW Factor VIII; modification; inhibitor activity; binding; antibody;  
KW von Willebrand factor; immune disorder.  
XX  
OS Synthetic.  
PN W09602572-A2.  
PD 01-FEB-1996.  
PF 14-JUL-1995; 95WO-BE00068.  
PR 14-JUL-1994; 94BE-0000666.  
PR (CROI-) CROIX ROUGE BELGIQUE.  
PA Di Giambattista M, Laub R;  
PI WPI; 1996-105861/11.  
PT Factor VIII antigenic polypeptide fragments and epitope(s) - also  
PT inhibitors of factor VIII and anti-inhibitors, useful for e.g.  
PT preventing and treating immune disorders involving inhibition of  
PT factor VIII binding  
XX  
PS Claim 8; Page 31; 45pp; French.

XX Peptides R90945-64 are derived from the factor VIII protein, esp. from a  
CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,  
CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII  
CC and derived peptides can be used to prevent the activity of inhibitors of  
CC factor VIII binding to von Willebrand factor, esp. antibodies, thus  
XX preventing or treating immune disorders.  
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 17; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOASRPYSF 9  
DB 1 nqasrpysf 9  
|||||

RESULT 2  
W32492  
ID W32492 standard; peptide; 8 AA.  
AC W32492;  
XX  
XX 21-APR-1998 (first entry)  
DT  
DE Helicostatin 3, which inhibits gut motility in the blowfly.  
DE  
KW Blowfly; callatostatin; insect neuropeptide; Leu-callatostatin;  
KW gut motility; cockroach; insecticide; blowfly; cydiastatin; helicostatin.  
XX  
XX Helicoverpa armigera.  
XX  
XX WO9735981-A1.  
PN  
XX  
XX 02-OCT-1997.  
PD  
XX  
XX 26-MAR-1997; 97WO-GR00843.  
PF  
XX  
XX 26-MAR-1996; 96GB-0006272.  
PR  
XX  
XX (QUEE-) QUEEN MARY & WESTFIELD COLLEGE.  
PA  
PI Duve H, East P, Johnsen AH, Thorpe A;  
PI  
XX WPI; 1997-489644/45.  
DR  
XX  
XX New callatostatin-like peptide(s) and DNA - are active as inhibitors  
PT of gut motility, used as insecticides, particularly against  
PT lepidopteran insects  
PT  
XX  
PS Claim 4; Page 62; 92pp; English.  
XX

The present sequence represents a specifically claimed callatostatin-like peptide which is active as an inhibitor of gut motility. The callatostatin-like peptides can be used as insecticides, particularly against lepidopteran insects. The invention relates to novel recombinant or isolated DNA sequences representing the Helicoverpa armigera helicostatin gene, the DraI fragment from Calliphora vomitoria or the Lucilia cuprina prothormone coding sequence. These sequences encode proteins of 228, 177 and 179 amino acids respectively.

XX  
SQ Sequence 8 AA;

Query Match 68.8%; Score 33; DB 18; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRPYSF 9  
|||||

Db 1 srpysf 6

RESULT 3  
P90351  
ID P90351 standard; protein; 7 AA.  
XX  
XX P90351;  
AC  
XX 01-NOV-1989 (first entry)  
DT  
XX Peptide fragment of human factor 8.  
DE  
XX Peptide fragments; human factor 8; Factor 8.  
KW  
XX Homo sapiens.  
OS  
XX EP321094-A.  
PN  
XX 21-JUN-1989.  
PD  
XX 16-NOV-1988; 88EP-0310801.  
PF  
XX 16-NOV-1987; 87US-0121217.  
PR  
XX (SCRI ) SCRIPPS CLINIC RES.  
PA  
XX Zimmerman TS, Fulcher CA;  
PI  
XX WPI; 1989-1779956/25.  
DR  
XX New polypeptide fragments  
PT - useful for neutralising activity of Factor 8 inhibitor.  
PT  
XX Claim 2; page 3 and page 6; 7pp; English.  
PS  
XX Peptide fragment of human factor 8, which  
CC neutralises the activity of the factor 8 inhibitor. Corresponds to  
CC residues 466-472 of factor 8 reported by 'Structure of human factor  
CC VIII' Nature Vol 312: 337-342 (1984).  
XX  
SQ Sequence 7 AA;

Query Match 64.6%; Score 31; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOASRP 6  
DB 2 nqasrp 7  
|||||

RESULT 4  
P90354  
ID P90354 standard; protein; 7 AA.  
XX  
XX P90354;  
AC  
XX 01-NOV-1989 (first entry)  
DT  
XX Peptide fragment of human Factor 8.  
DE  
XX Peptide fragments; human factor 8; Factor 8.  
KW  
XX Homo sapiens.  
OS  
XX EP321094-A.  
PN  
XX 21-JUN-1989.  
PD  
XX 16-NOV-1988; 88EP-0310801.  
PF  
XX

PR 16-NOV-1987; 87US-0121217.  
XX (SRI ) SCRIPPS CLINIC RES.  
XX  
PI Zimmerman TS, Fulcher CA;  
XX WPI; 1989-179956/25.  
XX New polypeptide fragments  
PT - useful for neutralising activity of Factor 8 inhibitor.  
XX  
PS Claim 2: page 3 and page 6; 7pp; English.  
XX  
CC Peptide fragment of human factor 8, which  
CC neutralises the activity of the factor 8 inhibitor. Corresponds to  
CC residues 1776-1782 of factor 8 reported by 'Structure of human factor  
CC VIII' Nature Vol 312: 337-342 (1984).  
XX  
SQ Sequence 7 AA;  
  
Query Match 64.6%; Score 31; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NQASRP 6  
DB 2 nqasrp 7  
  
RESULT 5  
W32493  
ID W32493 standard; peptide; 8 AA.  
XX  
AC W32493;  
XX  
DT 21-APR-1998 (first entry)  
XX  
DE Helicostatin 4, which inhibits gut motility in the blowfly.  
XX  
KW Blowfly; callatostatin; insect neuropeptide; Leu-callatostatin;  
KW gut motility; cockroach; insecticide; blowfly; cydiastatin; helicostatin.  
OS Helicoverpa armigera.  
XX  
PN W09735981-A1.  
XX  
PD 02-OCT-1997.  
XX  
PF 26-MAR-1997; 97WO-GB00843.  
XX  
PR 26-MAR-1996; 96GB-0006272.  
XX  
PA (QUEB-) QUEEN MARY & WESTFIELD COLLEGE.  
XX  
PI Duve H, East P, Johnson AH, Thorpe A;  
XX  
DR WPI; 1997-489644/45.  
XX  
PT New callatostatin-like peptide(s) and DNA - are active as inhibitors  
PT of gut motility, used as insecticides, particularly against  
PT lepidopteran insects  
XX  
PS Claim 4; Page 62; 92pp; English.  
XX  
CC The present sequence represents a specifically claimed callatostatin-  
CC like peptide which is active as an inhibitor of gut motility. The  
CC callatostatin-like peptides can be used as insecticides, particularly  
CC against lepidopteran insects. The invention relates to novel recombinant  
CC or isolated DNA sequences representing the Helicoverpa armigera  
CC heliocostatin gene, the Drai fragment from Calliphora vomitoria or the  
CC Lucilia cuprina prohormone coding sequence. These sequences encode  
CC proteins of 228, 177 and 179 amino acids respectively.

XX  
SQ Sequence 8 AA;  
  
Query Match 62.5%; Score 30; DB 18; Length 8;  
Best Local Similarity 83.3%; Pred. No. 2.1e+05;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 SRPYSF 9  
DB 1 arpsf 6  
  
RESULT 6  
R62417  
ID R62417 standard; Protein; 7 AA.  
XX  
AC R62417;  
XX  
DT 02-MAY-1995 (first entry)  
XX  
DE Dipteran neuropeptide, callatostatin 4.  
XX  
KW Neuropeptide; blowfly; C. vomitoria; callatostatin; inhibition;  
KW juvenile hormone; JH; insect development; metamorphosis;  
KW adult sexual maturity; reproduction; insect control;  
KW environmentally friendly.  
XX  
OS Calliphora vomitoria.  
XX  
PN W09420530-A.  
XX  
PD 15-SEP-1994.  
XX  
PF 14-MAR-1994; 94WO-GB00507.  
XX  
PR 12-MAR-1993; 93GB-0005120.  
XX  
PA (NAUY-) NAT UNIV HOSPITAL.  
PA (QUEB-) QUEEN MARY & WESTFIELD COLLEGE.  
PA (UTOR ) UNIV TORONTO.  
XX  
PI Duve H, Johnson A, Thorpe A, Tobe S;  
XX  
DR WPI; 1994-302962/37.  
XX  
PT Diptera neuropeptide(s), partic. from blowfly Calliphora  
PT vomitoria - are juvenile hormone prodn. Inhibitors, used for  
PT insect control, partic. cockroaches.  
XX  
PS Claim 6; Page 7; 11pp; English.  
XX  
CC The sequences given in R62414-18 are neuropeptides derived from the  
CC blowfly, Calliphora vomitoria. The peptides are extracted either  
CC from whole bodies or heads or C. vomitoria, or may be prepared by  
CC standard peptide synthesis. The peptides, called callatostatins,  
CC inhibit production of juvenile hormone (JH), which plays a crucial  
CC role in insect development by controlling metamorphosis, adult  
CC sexual maturity and reproduction. Callatostatin is the most powerful  
CC JH inhibitor known for cockroaches, with max. inhibition at 1 nM  
CC and ED50 at 0.1 nM. The peptides can be used for insect control  
CC without damage to the environment.  
XX  
SQ Sequence 7 AA;  
  
Query Match 60.4%; Score 29; DB 15; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 RPYSF 9  
DB 2 rpsf 6

```

XX PN WO9735981-A1.
XX PD 02-OCT-1997.
XX PF 26-MAR-1997; 97WO-GB00843.
XX PR 26-MAR-1996; 96GB-0006272.
XX PA (QUEE-) QUEEN MARY & WESTFIELD COLLEGE.
XX PI Duve H., East P, Johnsen AH, Thorpe A;
XX DR WPI; 1997-489644/45.
XX PT New callatostatin-like peptide(s) and DNA - are active as inhibitors
XX PT of gut motility, used as insecticides, particularly against
XX PS lepidopteran insects
XX PS Claim 25; Page 65; 92pp; English.
XX CC The present sequence represents a specifically claimed callatostatin-
XX CC like peptide which is active as an inhibitor of gut motility. The
XX CC callatostatin-like peptides can be used as insecticides, particularly
XX CC against lepidopteran insects. The invention relates to novel recombinant
XX CC or isolated DNA sequences representing the Helicoverpa armigera
XX CC helicostatin gene, the DraI fragment from Calliphora vomitoria or the
XX CC Lucilia cuprina prothomone coding sequence. These sequences encode
XX CC proteins of 228, 177 and 179 amino acids respectively.
XX SQ Sequence 8 AA;

Query Match 56.2%; Score 27; DB 18; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.1e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRPVSF 9
DB 1 arpynf 6

RESULT 9
Y47047
ID Y47047 standard; Peptide; 9 AA.
XX AC Y47047;
XX DT
XX DT 01-DEC-1999 (first entry)
XX DE Immunogenic peptide having a human leukocyte antigen binding motif #1658.
XX KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
XX KW immune response; T cell activation; major histocompatibility complex;
XX KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
XX KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
XX KW vaccine; immunisation.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9945954-A1.
XX PD 16-SEP-1999.
XX PF 13-MAR-1998; 98WO-US05039.
XX PR 13-MAR-1998; 98WO-US05039.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Kubo RT, Sidney J, Cells E, Grey HM, Southwood S;
XX XX

```

```

RESULT 7
W32499
ID W32499 standard; peptide; 7 AA.
XX AC W32499;
XX DT 21-APR-1998 (first entry)
XX DE Callatostatin 6, which inhibits gut motility in the blowfly.
XX KW Blowfly; callatostatin; insect neuropeptide; Leu-callatostatin;
XX KW gut motility; cockroach; insecticide; blowfly; cydiastatin; helicostatin.
XX OS Calliphora vomitoria.
XX PN WO9735981-A1.
XX PD 02-OCT-1997.
XX PF 26-MAR-1997; 97WO-GB00843.
XX PR 26-MAR-1996; 96GB-0006272.
XX PA (QUEE-) QUEEN MARY & WESTFIELD COLLEGE.
XX PI Duve H., East P, Johnsen AH, Thorpe A;
XX DR WPI; 1997-489644/45.
XX PT New callatostatin-like peptide(s) and DNA - are active as inhibitors
XX PT of gut motility, used as insecticides, particularly against
XX PS lepidopteran insects
XX PS Claim 16; Page 64; 92pp; English.
XX CC The present sequence represents a specifically claimed example of a
XX CC callatostatin peptide which is active as an inhibitor of gut motility.
XX CC The callatostatin peptides can be used as insecticides, particularly
XX CC against lepidopteran insects. The invention relates to novel recombinant
XX CC or isolated DNA sequences representing the Helicoverpa armigera
XX CC helicostatin gene, the DraI fragment from Calliphora vomitoria or the
XX CC Lucilia cuprina prothomone coding sequence. These sequences encode
XX CC proteins of 228, 177 and 179 amino acids respectively.
XX SQ Sequence 7 AA;

Query Match 60.4%; Score 29; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPYSF 9
DB 1 rpyssf 5

RESULT 8
W32506
ID W32506 standard; peptide; 8 AA.
XX AC W32506;
XX DT 21-APR-1998 (first entry)
XX DE Pteridostatin 2, which inhibits gut motility in the blowfly.
XX KW Blowfly; callatostatin; insect neuropeptide; Leu-callatostatin;
XX KW gut motility; cockroach; insecticide; blowfly; cydiastatin; helicostatin;
XX KW pteridostatin.
XX OS Pteronida salicis.

```

DR WPI; 1999-551214/46.  
 XX New immunogenic peptides with HLA binding motif, useful in treatment  
 PT and diagnosis of cancers and viral diseases  
 XX  
 XX Claim 1; Page 92; 150pp; English.  
 XX  
 CC Y45390 to Y48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful  
 CC therapeutically and for immunisation as above.  
 XX  
 XX Sequence 9 AA;

Query Match 56.2%; Score 27; DB 20; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 ASRPYSF 9  
 I : I I I  
 Db 3 adkpysy 9

RESULT 10  
 P90859  
 ID P90859 standard; protein; 7 AA.  
 XX  
 AC P90859;  
 XX  
 DT 29-JUN-1990 (first entry)  
 XX  
 DE New bradykinin analogue with D-Phe.  
 XX  
 KW Bradykinin analogue; bradykinin antagonist.

XX  
 FH Key Location/Qualifiers  
 FT Region 1..7  
 FT Misc-difference 5 /note="Claimed in claim 20"  
 FT /label=D-Phe  
 XX

PN W08901780-A.

XX 09-MAR-1989.

XX 29-AUG-1988; 88WO-US02959.

XX 02-SEP-1987; 87US-0092192.

XX (STEW ) STEWART JM.

XX Stewart JM, Vavrek RJ;

XX WPI; 1989-085400/11.

XX

PT New peptide bradykinin analogues -  
 PT with D-aminoacid in 7 position, useful as bradykinin antagonists  
 XX  
 PS Claim 20; Page 41; 54pp; English.  
 XX  
 CC It may be prep'd. by conventional liq. - or solid-phase peptide  
 CC synthesis methods. It is useful for treating local pain, inflammation  
 CC and swelling, rhinitis, hypotension, asthma, arthritis, diarrhoea,  
 CC irritable bowel syndrome, carcinoid syndrome, angina pain, and  
 CC anaphylactic or septic shock. Pharmaceutical compsns. can be made with  
 CC it.  
 XX  
 SQ Sequence 7 AA;

Query Match 52.1%; Score 25; DB 10; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPYSF 9  
 I I : I I  
 Db 1 rpfisf 5

RESULT 11  
 Y55007  
 ID Y55007 standard; peptide; 9 AA.

XX AC Y55007;

XX 18-FEB-2000 (first entry)

XX Mouse histocompatibility antigen binding peptide.

DE Active vaccination; autologous cell; transmembrane protein; cancer;  
 XX B cell Non-Hodgkin's lymphoma; B cell disease; autoimmune disorder;  
 KW Chronic lymphocytic leukaemia; B-cell regulatory disorder; therapy;  
 KW mouse; histocompatibility antigen.

XX OS Mus sp.

XX PN W09957981-A1.

XX 18-NOV-1999.

XX PF 07-MAY-1999; 99WO-US10065.

XX PR 08-MAY-1998; 98US-0084870.

XX (SLOK ) SLOAN KETTERING INST CANCER RES.

XX PI Agus DB, Scheinberg D, Roberts W, Zelenetz AD;

XX WPI; 2000-038920/03.

XX Novel compositions and methods utilizing a humoral immune response to  
 PT treat B cell Non-Hodgkin's lymphoma and other diseases .

XX Disclosure; Page 12; 43pp; English.

XX This sequence is a histocompatibility antigen binding peptide. The  
 CC invention relates to a method for active vaccination against autologous  
 CC cells expressing transmembrane proteins comprising administering to a  
 CC patient a vaccine composition comprising at least an immunogenic portion  
 CC of the extracellular domain of the transmembrane protein, or a xenogeneic  
 CC homolog thereof, coupled to or administered with a carrier protein  
 CC effective to break tolerance to the transmembrane protein and a  
 CC pharmaceutically acceptable adjuvant. The method of the invention is used  
 CC to treat a variety of cancers, especially B cell Non-Hodgkin's lymphoma  
 CC (NHL) and other B cell diseases such as chronic lymphocytic leukaemia,  
 CC autoimmune disorders and B-cell regulatory disorders. The method is also  
 CC useful for treating diseases and conditions where administration of  
 CC antibodies to transmembrane proteins has been shown to provide

CC therapeutic benefits, and especially where the target is also capable of  
 CC transducing or receiving a signal important for cell growth or function.  
 CC While therapeutic regimens using anti-CD20 concepts are potentially  
 CC effective for the treatment of B cell Non-Hodgkin's lymphoma, they are  
 CC passive therapies that do not directly involve the immune system of the  
 CC patient. Thus, these therapies may require continued administration of  
 CC the therapeutic agent for efficacy, and do not provide any long term  
 CC protection against recurrence. The present invention uses a humoral  
 CC immune response against CD20 found in B cells. The method does not need  
 CC long term administration of the therapeutic agent, and provides long term  
 CC protection.

XX Sequence 9 AA;

Query Match 52.1%; Score 25; DB 21; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QASRPY 7  
 I I I I  
 Db 3 qtskpy 8

RESULT 12

ID R71753 standard; peptide; 9 AA.

XX R71753;

XX 15-MAY-1995 (first entry)

XX Neurotensin receptor fluorescent probe.

XX Neurotensin; NT; receptor; probe; fluorescent.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-acylated with fluorescent label. See CC  
 FT below. Also this amino acid can be  
 FT substituted by Lys or Orn."

XX EP606804-A.

XX 20-JUL-1994.

XX 27-DEC-1993; 93EP-0403185.

XX 30-DEC-1992; 92CA-2086453.

XX (UYMC-) UNIV MCGILL.

XX Beaudet A, Faure M, Gaudreau P;

XX WPI; 1994-226757/28.

XX New fluorescent markers for neurotensin receptors - useful for in  
 PT vitro labelling of neurotensin receptors on cell surface and to  
 PT isolate neurotensin-receptor expressing cells

XX Claim 2; Page 5; 19pp; English.

XX The invention concerns highly sensitive fluorescent probes which  
 CC allow for rapid and precise characterisation of neurotensin  
 CC receptor binding properties on whole cells. The probes are of  
 CC formula R1-C(=X)-R in which R1 is a fluorophore chosen from  
 CC fluorescein, rhodamine, Blue fluorescent and Texas red; X is oxygen  
 CC or sulphur; and R is a neurotensin-type polypeptide sequence  
 CC containing the core sequence Arg-Arg-Pro-Tyr-Ile-Leu or  
 CC Arg-Arg-Pro-Trp-Ile-Leu. Specific sequences are given in R71745 -  
 CC R71772. The R1-CX- acyl group is linked to the peptide via the

CC N-terminus in position 1. The N-terminal amino acid may also be  
 CC substd. by Lys or Orn.  
 CC The present sequence represents one of the claimed peptide  
 CC sequences for R.

XX Sequence 9 AA;

Query Match 50.0%; Score 24; DB 15; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NOASRPY 7  
 I I I I  
 Db 1 nxprpy 7

RESULT 13

ID Y48827 standard; Peptide; 8 AA.

XX Y48827;

XX 10-DEC-1999 (first entry)

XX Membrane dipeptidase-binding retina homing peptide #13.

XX Homing peptide; organ; tissue; lung; pancreas; skin; retina; MDP;  
 KW prostate; ovary; lymph node; adrenal gland; liver; gut; tumour;  
 KW membrane dipeptidase.

XX Synthetic.

XX Homo sapiens.

XX WO9946284-A2.

XX 16-SEP-1999.

XX 10-MAR-1999; 99WO-US05284.

XX 13-MAR-1998; 98US-0042107.

XX 26-FEB-1999; 99US-0042107.

XX (BURN-) BURNHAM INST.

XX Rajotte D, Pasqualini R, Ruoslahti EI;

XX WPI; 1999-571717/48.

XX New peptides which selectively home to organs or tissues, used for,  
 PT e.g., identifying target ligands and for therapy of pathological  
 PT conditions

XX Example 6; Page 149; 193pp; English.

XX The present invention describes peptides that selectively home to a  
 CC tissue or organ. The peptides can be used for identifying an organ  
 CC or tissue, for identifying a target molecule expressed by an organ or  
 CC tissue or for treating an organ or tissue pathology, where the organ or  
 CC tissue is selected from prostate, lung, skin, retina, pancreas, gut,  
 CC ovary, adrenal gland, liver, and lymph node. The peptide bind to the  
 CC membrane dipeptidase (MDP). Y48618 to Y49066 represent sequences  
 CC which are used in the exemplification of the present invention.

XX Sequence 8 AA;

Query Match 47.9%; Score 23; DB 20; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRPY 7  
 I I I I



Db 3 srpy 6

RESULT 14  
R71755  
ID R71755 standard; peptide; 9 AA.  
XX AC  
XX R71755;  
XX  
XX 15-MAY-1995 (first entry)  
DT  
XX Neurotensin receptor fluorescent probe.  
DE  
XX Neurotensin; NT; receptor; probe; fluorescent.  
KW  
XX Synthetic.  
OS  
XX  
XX  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-acylated with fluorescent label. See CC  
FT below. Also this amino acid can be  
FT substituted by Lys or Orn."  
FT Modified-site 2 /label= Orn  
FT  
FT  
XX EP60804-A.  
XX  
XX 20-JUL-1994.  
PD  
XX  
XX 27-DEC-1993; 93EP-0403185.  
PF  
XX 30-DEC-1992; 92CA-2086453.  
PR  
XX (UYMC-) UNIV MCGILL.  
PA  
XX  
XX Beaudet A, Faure M, Gaudreau P;  
PI  
XX WPI; 1994-226757/28.  
DR  
XX New fluorescent markers for neurotensin receptors - useful for in  
PT vitro labelling of neurotensin receptors on cell surface and to  
PT isolate neurotensin-receptor expressing cells  
XX  
XX Claim 2; Page 5; 19pp; English.  
PS  
XX The invention concerns highly sensitive fluorescent probes which  
CC allow for rapid and precise characterisation of neurotensin  
CC receptor binding properties on whole cells. The probes are of  
CC formula R1-C(=X)-R in which R1 is a fluorophore chosen from  
CC fluorescein, rhodamine, Blue fluorescent and Texas red; X is oxygen  
CC or sulphur; and R is a neurotensin-type polypeptide sequence  
CC containing the core sequence Arg-Arg-bro-Tyr-Ile-Leu or  
CC Arg-Arg-Pro-Trp-Ile-Leu. Specific sequences are given in R71745 -  
CC R71772. The R1-CX- acyl group is linked to the peptide via the  
CC N-terminus in position 1. The N-terminal amino acid may also be  
CC substd. by Lys or Orn.  
CC The present sequence represents one of the claimed peptide  
CC sequences for R.  
XX  
XX Sequence 9 AA;

Query Match 45.8%; Score 22; DB 15; Length 9;  
Best Local Similarity 57.1%; Pred. NO. 2.1e+05;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NQASRPY 7  
| |||  
Db 1 nxprpy 7

RESULT 15  
Y55541

ID Y5541 standard; peptide; 9 AA.  
XX  
AC Y55541;  
XX  
DT 17-JAN-2000 (first entry)  
XX  
XX HLA binding plu-1 peptide.  
XX  
KW Human; cancer-associated polypeptide; plu-1; ovarian cancer; vaccine;  
KW breast cancer; immune response; cytotoxic T lymphocyte; imaging agent;  
KW therapeutic; plu-1 antigen; breast tumor; tumor antigen; epitope.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
XX WO9949034-A1.  
PN  
XX 30-SEP-1999.  
PD  
XX 19-MAR-1999; 99WO-GB00866.  
PF  
XX 20-MAR-1998; 98GB-0005877.  
PR  
XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.  
PA  
XX Taylor-papadimitriou J;  
PI  
XX WPI; 1999-591090/50.  
DR  
XX  
XX New nucleic acid encoding the cancer-associated polypeptide plu-1, for  
PT diagnosis, treatment and prevention of cancer, especially of breast and  
PT ovary -  
XX  
PS Example 2; Fig 12; 173pp; English.  
XX  
XX The invention relates to a human cancer-associated polypeptide plu-1. The  
CC plu-1 polypeptide can be recombinantly expressed by standard recombinant  
CC methodology. Detection of the plu-1 nucleic acid or the polypeptide is  
CC used for the following: (i) diagnosis (including imaging) and prognosis  
CC of, and determination of susceptibility to, cancer, specifically ovarian  
CC or breast cancer; and (ii) treating cancer (by inducing an immune response  
CC against cancer cells, e.g. as a vaccine, or by antisense inhibition).  
CC Antigens derived from the polypeptide are used to generate activated  
CC cytotoxic T lymphocytes, or dendritic cells, for subsequent return to  
CC the patient for treatment of cancer. The polypeptide may also be used to  
CC identify inhibitors of plu-1 activity. Fragments of the polypeptide, and  
CC antibodies raised against plu-1, are useful as assay and imaging agents,  
CC also therapeutically (to induce an anti-idiotypic response or where  
CC conjugated to cytotoxic agents). The plu-1 antigen is expressed more  
CC commonly in breast tumors than some known tumor antigens. Sequences  
CC Y55320-629 represent predicted peptides from the plu-1 polypeptide which  
CC may bind to the human class I alleles B27, A2, A3 and A11.  
XX  
XX Sequence 9 AA;

Query Match 45.8%; Score 22; DB 20; Length 9;  
Best Local Similarity 57.1%; Pred. NO. 2.1e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QASRPYS 8  
|:|:|  
Db 1 gaardyt 7

Search completed: January 17, 2001, 13:42:21  
Job time: 135 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:41:26 ; Search time 58.21 Seconds  
(without alignments)  
18.122 Million cell updates/sec

Title: US-08-765-837-4  
Perfect score: 48  
Sequence: 1 NQASRPYSF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 467

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	33.3	7	10	P93233 lycopersico
2	15	31.2	7	8	O98866 spinacia ol
3	15	31.2	9	12	Q67605 squash leaf
4	15	31.2	9	12	Q67606 squash leaf
5	14	29.2	7	11	Q63668 rattus norv
6	14	29.2	8	2	Q9R7T2 escherichia
7	14	29.2	8	5	Q9VVJ2 drosophila
8	14	29.2	8	5	Q9TWH6 perineiris
9	14	29.2	8	6	Q9TRX8 bos taurus
10	14	29.2	9	1	Q50832 methanococ
11	14	29.2	9	2	Q53914 streptomyce
12	14	29.2	9	4	Q9UKJ6 homo sapien
13	14	29.2	9	6	Q9TRU7 bos taurus
14	13	27.1	7	12	Q66113 cherry leaf
15	13	27.1	8	4	Q9UJ50 homo sapien
16	13	27.1	8	8	Q9T4Y2 asterina pe
17	13	27.1	8	11	Q9QVD3 rattus sp.
18	13	27.1	8	11	P82598 rattus norv
19	13	27.1	8	12	O89965 polyomaviru

20	13	27.1	9	2	Q45852
21	13	27.1	9	4	O16605
22	13	27.1	9	4	Q9UC36
23	13	27.1	9	6	Q28112
24	13	27.1	9	8	P92072
25	13	27.1	9	8	Q9T2L1
26	13	27.1	9	8	Q9T2L0
27	13	27.1	9	11	O88889
28	13	27.1	9	12	O69473
29	13	27.1	9	12	Q9PYK1
30	13	27.1	9	12	Q9IBM8
31	12	25.0	8	4	Q15898
32	12	25.0	8	5	P82618
33	12	25.0	8	6	Q28866
34	12	25.0	8	11	Q9TLD7
35	12	25.0	9	4	Q15999
36	12	25.0	9	4	Q9UMF3
37	12	25.0	9	4	Q9UM87
38	12	25.0	9	4	Q9UGE4
39	12	25.0	9	5	Q9TWV0
40	12	25.0	9	11	P97889
41	12	25.0	9	12	Q82622
42	12	25.0	9	13	Q9PRJ4
43	12	25.0	9	13	Q9IAI4
44	11	22.9	7	12	Q65578
45	11	22.9	8	7	Q95213

#### ALIGNMENTS

RESULT 1  
P93233 PRELIMINARY; PRT; 7 AA.  
AC P93233;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE (EC 4.4.1.14)  
DE (ACC SYNTHASE) (FRAGMENT).  
GN LE-ACS1B.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasi...ds 1;  
OC Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97351561; PubMed=9207843;  
RA Oetliker J.H., Olson D.C., Shiu O.Y., Yang S.F.;  
RT "Differential induction of seven l-aminocyclopropane-1-carboxylate  
RT synthase genes by elicitor in suspension cultures of tomato  
RT (Lycopersicon esculentum).";  
RL Plant Mol. Biol. 34:275-286(1997).  
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE = l-  
CC -1- AMINOCYCLOPROPANE-1-CARBOXYLATE + METHYLTHIOADENOSINE.  
CC -1- COFACTOR: PYRIDOXAL-PHOSPHATE.  
DR EMBL; U75692; AAC49682.1; -.  
KW Lyase.  
FT NON\_TER 1 1  
SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;

Query Match 33.3%; Score 16; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRP 6  
DB 1 SRP 3

RESULT 2

```

098866
ID O98866 PRELIMINARY; PRT; 7 AA.
AC O98866;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE CYTOCHROME B/F SUBUNIT IV (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Chloroplast.
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta;
OC Magnoliophyta: eudicotyledons: core eudicots: Caryophyllidae;
OC Caryophyllales: Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86120353; PubMed=3003688;
RA Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
RT "Spinach plastid genes coding for initiation factor IF-1, ribosomal
protein S11 and RNA polymerase alpha-subunit.";
RL Nucleic Acids Res. 14:1029-1044(1986).
DR EMBL; X03496; CAA27215.1; -.
KW Chloroplast.
FT NON_TER 1 1
SQ SEQUENCE 7 AA; 907 MW; 644729D77409C420 CRC64;

Query Match 31.2%; Score 15; DB 8; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPY 7
II:
DB 3 RPF 5

RESULT 3
Q67605
ID Q67605 PRELIMINARY; PRT; 9 AA.
AC Q67605;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TReMBLrel. 08, Last annotation update)
DE A COMPONENT DNA (FRAGMENT).
GN AL1.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082450; PubMed=1984669;
RA Lazarowitz S.G.;
RT "Molecular characterization of two bipartite geminiviruses causing
squash leaf curl disease: role of viral replication and movement
functions in determining host range.";
RL Virology 180:70-80(1991).
DR EMBL; M63155; AAA47823.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1118 MW; 2B30D5B457645417 CRC64;

Query Match 31.2%; Score 15; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 3.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9
II:
DB 5 PNSF 8

RESULT 4
Q67606
ID Q67606 PRELIMINARY; PRT; 9 AA.
AC Q67606;

```

```

DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TReMBLrel. 08, Last annotation update)
DE A COMPONENT DNA (FRAGMENT).
GN ALL.
OS Squash leaf curl virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10829;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91082450; PubMed=1984669;
RA Lazarowitz S.G.;
RT "Molecular characterization of two bipartite geminiviruses causing
squash leaf curl disease: role of viral replication and movement
functions in determining host range.";
RL Virology 180:70-80(1991).
DR EMBL; M63155; AAA47823.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1118 MW; 2B30D5B457645417 CRC64;

Query Match 31.2%; Score 15; DB 12; Length 9;
Best Local Similarity 75.0%; Pred. No. 3.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9
II:
DB 5 PNSF 8

RESULT 5
Q63668
ID Q63668 PRELIMINARY; PRT; 7 AA.
AC Q63668;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TReMBLrel. 08, Last annotation update)
DE VASOPRESSIN V2 RECEPTOR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=95396550; PubMed=7667072;
RA Mandon B., Bellanger A.C., Elalouf J.M.;
RT "Inverse PCR-mediated cloning of the promoter for the rat vasopressin
V2 receptor gene.";
RL Pflügers Arch. 430:12-18(1995).
DR EMBL; X83264; CAA58237.1; -.
SQ SEQUENCE 7 AA; 703 MW; 75A767287DC6D6F0 CRC64;

Query Match 29.2%; Score 14; DB 11; Length 7;
Best Local Similarity 60.0%; Pred. No. 3.7e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 QASRP 6
II:
DB 2 QAGLP 6

RESULT 6
Q9R7T2
ID Q9R7T2 PRELIMINARY; PRT; 8 AA.
AC Q9R7T2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE HYPOTHETICAL PROTEIN HI0004 (FRAGMENT).
GN YOFG.
OS Escherichia coli.

```

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,  
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-Kb DNA Sequence of Escherichia coli K-12 Genome Corresponding  
 RT to the 12.7-28.0 min Region on the Linkage Map.";  
 RL DNA Res. 3:137-155(1996).  
 DR EMBL; D90705; BAA35310.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 964 MW; DF133B1DD04B476A CRC64;

Query Match 29.2%; Score 14; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PY 7  
 DB 2 PY 3

RESULT 7  
 ID Q9VJ2 PRELIMINARY; PRT; 8 AA.  
 AC Q9VJ2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE CG6143 PROTEIN (FRAGMENT).  
 DE CG6143.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyroidae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Purl V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.G.,  
 RA Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2193(2000).  
 DR EMBL; AE003523; AAF49319.1; -;  
 DR FLYBASE; FBgn0036743; CG6143.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 919 MW; 740045A2C866CB47 CRC64;

Query Match 29.2%; Score 14; DB 5; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PY 7  
 DB 1 PY 2

RESULT 8  
 ID Q9TWH6 PRELIMINARY; PRT; 8 AA.  
 AC Q9TWH6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE BIOACTIVE PEPTIDE P4-PUTATIVE ESOPHAGEAL NEUROREGULATOR.  
 OS Perinereis vancaurica.  
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;  
 OC Phyllodocta; Nereididae; Perinereis.  
 OX NCBI\_TaxID=6355;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95323338; PubMed=7599979;  
 RA Takahashi T., Furukawa Y., Muneoka Y., Matsushima O., Ikeda T.,  
 RA Fujita T., Minakata H., Nomoto K.;  
 RT "Isolation and characterization of four novel bioactive peptides from  
 RT a polychaete annelid, Perinereis vancaurica.";  
 RL Comp. Biochem. Physiol. C,  
 RL Pharmacol. Toxicol. Endocrinol. 110:297-304(1995).  
 SQ SEQUENCE 8 AA; 989 MW; 954772CAA87B0B59 CRC64;

Query Match 29.2%; Score 14; DB 5; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PY 7  
 DB 7 PY 8

RESULT 9  
 ID Q9TRX8 PRELIMINARY; PRT; 8 AA.  
 AC Q9TRX8;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE OSTEOPONTIN.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;

RN SEQUENCE.  
 RP MEDLINE=91282766; PubMed=1676261;  
 RA Prince C.W., Dickie D., Krumdieck C.L.;  
 RT "Osteopontin, a substrate for transglutaminase and factor XIII  
 activity";  
 RL Biochem. Biophys. Res. Commun. 177:1205-1210(1991).  
 SQ SEQUENCE 8 AA; 920 MW; 05DAFAF76322D767 CRC64;

Query Match 29.2%; Score 14; DB 6; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYSF 9  
 :|:|  
 DB 4 KPXYF 8

RESULT 10  
 Q50832 PRELIMINARY; PRT; 9 AA.  
 AC Q50832;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DE INTERGENIC AT-RICH DNA SEQUENCE (FRAGMENT).  
 OS Methanococcus voltae.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 OX NCBI\_TaxID=2188;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85230552; PubMed=4006907;  
 RA Bollschweiler C., Kuehn R., Klein A.;  
 RT "Non-repetitive AT-rich sequences are found in intergenic regions of  
 Methanococcus voltae DNA";  
 RL EMBO J. 4:805-809(1985).  
 DR EMBL; X02518; CAA26355.1; -.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1087 MW; 99ED005DC4040A5A CRC64;

Query Match 29.2%; Score 14; DB 1; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 1 NQASR 5  
 :|:|  
 DB 4 NRAIR 8

RESULT 11  
 Q53914 PRELIMINARY; PRT; 9 AA.  
 AC Q53914;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DE CURD PROTEIN (FRAGMENT).  
 OS Streptomyces cyanus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1904;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bergh S.T., Uhlen M.;  
 RL Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M33704; AAA26724.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 1027 MW; 995BDDDDC4140AB1 CRC64;

Query Match 29.2%; Score 14; DB 2; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 QASRP 6  
 :|:|  
 DB 5 RAAQP 9

RESULT 12  
 Q9UKJ6 PRELIMINARY; PRT; 9 AA.  
 AC Q9UKJ6;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DE ANDROGEN RECEPTOR (FRAGMENT).  
 GN AR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang W., Li X.Q., Wu Q.F.;  
 RT "A splice-site mutation in Androgen Receptor gene";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF159087; AAF04001.1; -.  
 KW Receptor.  
 FT NON\_TER 1  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1272 MW; 6F2B8415B331E684 CRC64;

Query Match 29.2%; Score 14; DB 4; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 SRPY 7  
 :|:|  
 DB 6 SRMY 9

RESULT 13  
 Q9TRU7 PRELIMINARY; PRT; 9 AA.  
 AC Q9TRU7;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DE GAP-3, GTPASE-ACTIVATING PROTEIN.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE  
 RX MEDLINE=92112868; PubMed=1309786;  
 RA Nice E.C., Fabri L., Hammacher A., Hoiden J., Simpson R.J.,  
 RA Burgess A.W.;  
 RT "The purification of a Rap1 GTPase-activating protein from bovine  
 brain cytosol";  
 RL J. Biol. Chem. 267:1546-1553(1992).  
 SQ SEQUENCE 9 AA; 1063 MW; 89EDA77B47604B5A CRC64;

Query Match 29.2%; Score 14; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PY 7  
 :|:|

Db 5 PY 6

## RESULT 14

Q66113  
ID Q66113 PRELIMINARY; PRT; 7 AA.  
AC Q66113;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE C-TERMINUS OF THE VIRAL REPLICASE (FRAGMENT).  
OS cherry leaf roll virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;  
OC Nepovirus.  
OX NCBI\_TaxID=12615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WALNUT;  
RA Borja M.;  
RL Thesis (1992), Biologia Molecular y Virologia Vegetal, CIT-INIA.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WALNUT;  
RX MEDLINE=96124520; PubMed=8560786;  
RA Borja M., Sanchez F., Rowhani A., Bruening G., Ponz F.;  
RT "Long, nearly identical untranslated sequences at the 3' terminal  
RT regions of the genomic RNAs of cherry leafroll virus (walnut  
RT strain).";  
RL Virus Genes 10:245-252(1995).  
DR EMBL; Z34265; CAA84019.1; -.  
KW Repeat.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 800 MW; 7417672BBC6D740 CRC64;

## Query Match

Best Local Similarity 27.1%; Score 13; DB 12; Length 7;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 QASRP 6  
DB 2 QACLP 6

## RESULT 15

Q90J50  
ID Q90J50 PRELIMINARY; PRT; 8 AA.  
AC Q90J50;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE LATROPHILIN-2 (FRAGMENT).  
GN LPHH1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA White G.R.M., Varley J.M., Heighway J.;  
RT "Isolation and characterisation of a human homologue of the  
RT latrophilin gene from a region of lp31.1 implicated in breast  
RT cancer.";  
RL Oncogene 17:3513-3519(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA White G.R.M., Varley J.M., Heighway J.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ244509; CAB60204.1; -.  
FT NON\_TER 1  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 1025 MW; 4D59D763344AAB0B CRC64;

Query Match 27.1%; Score 13; DB 4; Length 8;  
Best Local Similarity 25.0%; Pred. No. 3.7e+05;  
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRPY 7  
DB 4 NKPF 7

Search completed: January 17, 2001, 13:44:40  
Job time: 194 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:41:01 ; Search time 36.59 Seconds  
(without alignments)  
16.701 Million cell updates/sec

Title: US-08-765-837-4  
Perfect score: 48  
Sequence: 1 NQASRPYSF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 787

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	60.4	8	D47393	neuropeptide calla
2	20	41.7	8	E47393	neuropeptide calla
3	17	35.4	9	S70332	endosperm protein,
4	16	33.3	9	PC7076	spectrin alpha cha
5	15	31.2	6	I51317	bHLH transcription
6	15	31.2	7	E33932	Ig mu chain D regi
7	15	31.2	8	PT0030	inulinase (EC 3.2.
8	15	31.2	8	PH0934	T-cell receptor be
9	15	31.2	9	A61057	Thr-6 bradykinin -
10	15	31.2	9	A26744	bradykinin-like pe
11	15	31.2	9	A61363	bradykinin - como
12	15	31.2	9	A60579	bradykinin-like pe
13	15	31.2	9	B60246	ornitho-kinin - ch
14	15	31.2	9	S65433	bradykinin - horn
15	15	31.2	9	S78420	ribosomal protein
16	15	31.2	9	A43065	hydroxyproline-3-b
17	15	31.2	9	S15594	orf 1 rara 5'-regi
18	14	29.2	5	JS0319	subesophageal gang
19	14	29.2	5	PT0714	T-cell receptor be
20	14	29.2	6	PT0618	T-cell receptor be
21	14	29.2	6	PT0715	T-cell receptor be
22	14	29.2	7	A61324	dermorphin - Rohde
23	14	29.2	7	A60139	fatty-acid synthas
24	14	29.2	7	S36662	dermorphin (lys-7)
25	14	29.2	7	S21230	dermorphin (trp-4,
26	14	29.2	8	S21288	cardioacceleratory
27	14	29.2	8	S66646	serum albumin - do
28	14	29.2	8	B45800	xenopsin-related p
29	14	29.2	9	JS0302	

## ALIGNMENTS

## RESULT 1

D47393  
neuropeptide callatostatin 4 - bluebottle fly (Calliphora vomitoria)  
C:Species: Calliphora vomitoria  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 28-Apr-1995  
C:Accession: D47393  
R:Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993  
A:Title: Callatostatin: neuropeptides from the blowfly Calliphora vomitoria with seq  
A:Reference number: A47393; MUID:93211980  
A:Accession: D47393  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <DUV>  
A:Experimental source: thoracic ganglia  
A>Note: sequence extracted from NCBI backbone (NCBIP:128479)

Query Match 60.4%; Score 29; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPYSF 9  
DB 2 RPYSF 6

## RESULT 2

E47393  
neuropeptide callatostatin 5 - bluebottle fly (Calliphora vomitoria)  
C:Species: Calliphora vomitoria  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 03-Mar-1995  
C:Accession: E47393  
R:Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993  
A:Title: Callatostatin: neuropeptides from the blowfly Calliphora vomitoria with seq  
A:Reference number: A47393; MUID:93211980  
A:Accession: E47393  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <DUV>  
A:Experimental source: whole flies  
A>Note: sequence extracted from NCBI backbone (NCBIP:128482)

Query Match 41.7%; Score 20; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RPYSF 9  
DB 3 PYDF 6

30 14 29.2 9 2 A60320 xenopsin related p  
31 14 29.2 9 2 S30494 cat gene leader pe  
32 14 29.2 9 2 B24362 chloramphenicol O-  
33 14 28.2 9 2 D48186 ATPase K1 subunit  
34 14 29.2 9 2 S77984 cytochrome-c oxida  
35 14 29.2 9 2 S10784 enamelin I - bovin  
36 14 29.2 9 2 PH1591 Ig H chain V-D-J I  
37 13 27.1 4 2 PT0675 T-cell receptor be  
38 13 27.1 5 2 PT0525 T-cell receptor be  
39 13 27.1 6 2 A61140 sperm acrosomal pr  
40 13 27.1 6 2 PT0518 T-cell receptor be  
41 13 27.1 6 2 PT0662 T-cell receptor be  
42 13 27.1 6 4 A35039 hypothetical colla  
43 13 27.1 7 2 PT0515 T-cell receptor be  
44 13 27.1 7 2 PT0671 T-cell receptor be  
45 13 27.1 8 2 PT0559 T-cell receptor be

```

RESULT 3
S70332
endosperm protein, 10K - rye (fragment)
C:Species: Secale cereale (rye)
C>Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C:Accession: S70332
R:Roche, A.; Calero, M.; Soriano, F.; Mendez, E.
Biochim. Biophys. Acta 1295, 13-22, 1996
A:Title: Identification of major rye secalins as coeliac immunoreactive proteins.
A:Reference number: S70327; MUID:96283789
A:Accession: S70332
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <ROC>

Query Match 35.4%; Score 17; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9
   | |
Db 3 PYEY 6

RESULT 4
PC7076
spectrin alpha chain, non-erythroid - mouse (fragment)
N:Alternate names: fodrin alpha chain
C:Species: Mus musculus (house mouse)
C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: PC7076
R:Tsuigita, A.; Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Matsui, T.; Watanabe, Y.;
Electrophoresis 21, 1853-1871, 2000
A:Title: Proteome analysis of mouse brain: Two-dimensional electrophoresis profiles of
A:Reference number: PC7072
A:Accession: PC7076
A:Molecule type: protein
A:Residues: 1-9 <TSU>
A:Experimental source: strain C57BL/6Cr Slc, male; brain, striatum
C:Keywords: brain

Query Match 33.3%; Score 16; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NSASR 5
   | |
Db 1 NSATR 5

RESULT 5
I51317
bHLH transcription factor inhibitor - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I51317
R:Zhang, H.; Reynaud, S.; Kloc, M.; Etkin, L.D.; Spohr, G.
Mech. Dev. 50, 119-130, 1995
A:Title: Id gene activity during Xenopus embryogenesis.
A:Reference number: I51316; MUID:95344988
A:Accession: I51317
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-6 <ZHA>
A:Cross-references: GB:S79038; NID:g1042006; PIDN:AA014294.1; PID:g4261994
C:Genetics:
A:Gene: Xidib

Query Match 31.2%; Score 15; DB 2; Length 6;
Best Local Similarity 40.0%; Pred. No. 2e+05;
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 ASRPY 7
   | |
Db 2 ATEPF 6

RESULT 6
E33932
Ig mu chain D region (E7) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 16-Aug-1996
C:Accession: E33932
R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989
A:Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated ger
A:Reference number: A33932; MUID:89282823
A:Accession: E33932
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-7 <BAC>
A:Cross-references: GB:M27106
C:Keywords: immunoglobulin

Query Match 31.2%; Score 15; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPY 7
   | |
Db 1 RPF 3

RESULT 7
PT0030
inulinase (EC 3.2.1.7) - Aspergillus ficuum (fragment)
N:Alternate names: inulase
C:Species: Aspergillus ficuum
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 07-May-1999
C:Accession: PT0030
R:Etalibi, M.; Barattii, J.C.
Agric. Biol. Chem. 54, 61-68, 1990
A:Title: Molecular and kinetic properties of Aspergillus ficuum inulinases.
A:Reference number: PT0030; MUID:90344234
A:Accession: PT0030
A:Molecule type: protein
A:Residues: 1-8 <ETT>
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 31.2%; Score 15; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPY 7
   | |
Db 5 QPY 7

RESULT 8
PH0934
T-cell receptor beta chain V-D-J region (clone 5) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C:Accession: PH0934
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental alle
A:Reference number: PH0891; MUID:92078857
A:Accession: PH0934

```

A:Molecule type: mRNA

A:Residues: 1-8 <GOL>

A:Experimental notes: complete Freund's adjuvant-immunized lymph node

A>Note: the authors translated the codon CAG for residue 7 as Glu

C:Keywords: T-cell receptor

Query Match 31.2%; Score 15; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOA 3

Db 6 NOA 8

RESULT 9

A61057

Thr-6 bradykinin - scoliid wasp (Colpa interrupta)

C:Species: Colpa interrupta

C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 18-Aug-2000

C:Accession: A61057

R:Piek, T.; Hue, B.; Mantel, P.; Nakajima, T.; Pelhate, M.; Yasuhara, T.

Comp. Biochem. Physiol. C 96, 157-162, 1990

A:Title: Threonine(6)-bradykinin in the venom of the wasp Colpa interrupta (F.) presynap

A:Reference number: A61057; MUID:91130217

A:Accession: A61057

A:Molecule type: protein

A:Residues: 1-9 <PIE>

C:Superfamily: unassigned animal peptides

C:Keywords: bradykinin; presynaptic neurotoxin; venom

Query Match 31.2%; Score 15; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYSF 9

Db 1 RPPGF 5

RESULT 10

A26744

bradykinin-like peptide - garden dagger wasp

N:Alternate names: Thr-6-bradykinin

C:Species: Megascolia flavifrons (garden dagger wasp)

C>Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 18-Aug-2000

C:Accession: A26744

R:Yasuhara, T.; Mantel, P.; Nakajima, T.; Piek, T.

Toxicon 25, 527-535, 1987

A:Title: Two kinins isolated from an extract of the venom reservoirs of the solitary wasp

A:Reference number: A94322; MUID:87293024

A:Accession: A26744

A:Molecule type: protein

A:Residues: 1-9 <YAS>

C:Superfamily: unassigned animal peptides

Query Match 31.2%; Score 15; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYSF 9

Db 1 RPPGF 5

RESULT 11

A61363

bradykinin - common frog

C:Species: Rana temporaria (common frog)

C>Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 18-Aug-2000

C:Accession: A61363

R:Anastasi, A.; Erspamer, V.; Bertaccini, G.

Comp. Biochem. Physiol. A 14, 43-52, 1965

A:Title: Occurrence of bradykinin in the skin of Rana temporaria.

A:Reference number: A61363

A:Accession: A61363

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <ANA>

C:Superfamily: unassigned animal peptides

C:Keywords: skin

Query Match 31.2%; Score 15; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYSF 9

Db 1 RPPGF 5

RESULT 12

A60579

bradykinin-like peptide - slider turtle

C:Species: Pseudemys scripta (slider)

C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 18-Aug-2000

C:Accession: A60579

R:Conlon, J.M.; Hicks, J.W.; Smith, D.D.

Endocrinology 126, 985-991, 1990

A:Title: Isolation and biological activity of a novel kinin ([Thr(6)]bradykinin) from

A:Reference number: A60579; MUID:90126625

A:Accession: A60579

A:Molecule type: protein

A:Residues: 1-9 <CON>

C:Comment: This peptide increases aortic blood flow but, unlike bradykinin in mammal

C:Superfamily: unassigned animal peptides

C:Keywords: plasma

Query Match 31.2%; Score 15; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYSF 9

Db 1 RPPGF 5

RESULT 13

B60246

ornitho-kinin - chicken

C:Species: Gallus gallus (chicken)

C>Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 18-Aug-2000

C:Accession: B60246

R:Kimura, M.; Sueyoshi, T.; Morita, T.; Tanaka, K.; Iwanaga, S.

Adv. Exp. Med. Biol. 247A, 359-367, 1989

A:Title: Ornitho-kininogen and ornitho-kinin: isolation, characterization and chemica

A:Reference number: A60246; MUID:90102072

A:Accession: B60246

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <KIM>

C:Superfamily: unassigned animal peptides

Query Match 31.2%; Score 15; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYSF 9

Db 1 RPPGF 5

## RESULT 14

S65433  
bradykinin - horn fly (fragment)  
C:Species: Haematobia irritans (horn fly)  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997  
C:Accession: S65433  
R:Wijffels, G.; Fitzgerald, C.; Gough, J.; Riding, G.; Elvin, C.; Kemp, D.; Willadsen, P.  
Eur. J. Biochem. 237, 414-423, 1996  
A:Title: Cloning and characterisation of angiotensin-converting enzyme from the dipteran  
A:Reference number: S65431; MUID:96215437  
A:Accession: S65433  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <WL>  
A>Note: the source is designated as Haematobia irritans exigua

Query Match 31.2%; Score 15; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 RPYSF 9  
|||  
Db 1 RPPGF 5

## RESULT 15

S78420  
ribosomal protein RL41, mitochondrial [validated] - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jul-2000  
C:Accession: S78420  
R:Goldschmidt-Reisin, S.; Graack, H.R.  
submitted to the Protein Sequence Database, February 1998  
A:Reference number: S78411  
A:Accession: S78420  
A:Molecule type: protein  
A:Residues: 1-9 <GOL>  
A>Note: the protein is designated as mitochondrial ribosomal protein L41  
C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 31.2%; Score 15; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NQASR 5  
|||  
Db 2 NQXHR 6

Search completed: January 17, 2001, 13:43:36  
Job time: 155 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:42:31 ; Search time 20.9 seconds  
(without alignments)  
13.907 Million cell updates/sec

Title: US-08-765-837-4  
Perfect score: 48  
Sequence: 1 NQASRPYSF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 213

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	33	68.8	8	1 ALL3_CVDPO	P82155 cydia pomon
2	30	62.5	8	1 ALL4_CVDPO	P82155 cydia pomon
3	29	60.4	8	1 ALL4_CALVO	P41840 calliphora
4	25	52.1	9	1 ALL10_CARMA	P81813 carcinus ma
5	24	50.0	8	1 ALL15_CARMA	P81818 carcinus ma
6	21	43.8	7	1 ALL3_CARMA	P81806 carcinus ma
7	21	43.8	7	1 ALL4_CARMA	P81807 carcinus ma
8	21	43.8	7	1 ALL5_CARMA	P81808 carcinus ma
9	21	43.8	8	1 ALL6_CARMA	P81819 carcinus ma
10	21	43.8	8	1 ALL7_CARMA	P81809 carcinus ma
11	21	43.8	8	1 ALL8_CARMA	P81811 carcinus ma
12	21	43.8	8	1 ALL9_CARMA	P81812 carcinus ma
13	20	41.7	8	1 ALL5_CALVO	P41841 calliphora
14	19	39.6	9	1 FLA2_TREHY	P80159 treponema h
15	18	37.5	8	1 ALL7_CARMA	P81820 carcinus ma
16	17	35.4	5	1 ALL4_CARMA	P81817 carcinus ma
17	17	35.4	8	1 ALL8_CARMA	P81821 carcinus ma
18	17	35.4	8	1 ALL5_CVDPO	P82156 cydia pomon
19	16	33.3	9	1 ALL1_CARMA	P81814 carcinus ma
20	16	33.3	9	1 FAR9_ASCSU	P43172 ascheris suu
21	14	29.2	5	1 SUGA_ACHDO	P19991 acheta dome
22	14	29.2	7	1 ALL2_CARMA	P81805 carcinus ma
23	14	29.2	7	1 GFPP_MOUSE	P99025 mus musculus
24	14	29.2	8	1 ALL2_CARMA	P81815 carcinus ma
25	14	29.2	8	1 ALL1_CVDPO	P82152 cydia pomon
26	14	29.2	8	1 ALL6_CVDPO	P82157 cydia pomon
27	14	29.2	8	1 B44K_PORGI	P81886 porphyromon
28	14	29.2	9	1 COXE_THUOB	P80975 thunnus obe
29	14	29.2	9	1 LPCA_STAAU	P36884 staphylococ
30	14	29.2	9	1 NEUX_HUMAN	P04277 homo sapien
31	14	29.2	9	1 NEUX_RAT	P11382 rattus norv
32	14	29.2	9	1 UPA7_HUMAN	P30093 homo sapien
33	13	27.1	7	1 ALL7_CVDPO	P82158 cydia pomon

RESULT 1  
ALL3\_CVDPO  
ID ALL3\_CVDPO STANDARD; PRT; 8 AA.  
AC P82154;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE CYDIATOSTATIN 3.  
OS Cydia pomonella (Codling moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=LARVA;  
RX MEDLINE=98054539; PubMed=9392829;  
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,  
RA Davey M., East P.D., Thorpe A.;  
RT "Lepidopteran peptides of the allatostatin superfamily.";  
RL Peptides 18:1301-1309(1997).  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 8  
SQ SEQUENCE 8 AA; 926 MW; C82879D5AB477415 CRC64;

Query Match 68.8%; Score 33; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 8.9e+04; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;

QY 4 SRPYSF 9  
| | | | |  
Db 1 SRPYSF 6

RESULT 2  
ALL4\_CVDPO  
ID ALL4\_CVDPO STANDARD; PRT; 8 AA.  
AC P82155;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE CYDIATOSTATIN 4.  
OS Cydia pomonella (Codling moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=LARVA;  
RX MEDLINE=98054539; PubMed=9392829;  
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,  
RA Davey M., East P.D., Thorpe A.;  
RT "Lepidopteran peptides of the allatostatin superfamily.";  
RL Peptides 18:1301-1309(1997).  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

O68928 salmonella  
O68936 serratus  
P56576 rattus norv  
P19990 leucoplaca  
P15507 bos taurus  
P34966 cavla porce  
O54296 salmonella  
P41517 calliphora  
P42985 lepidoptera  
P41866 calliphora  
P25423 melolontha  
P38644 mus musculus

```

KW Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 910 MW; 922879D5AB47740D CRC64;

Query Match 62.5%; Score 30; DB 1; Length 8;
Best Local Similarity 83.3%; Pred. No. 8.9e+04;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRPYSF 9
Db 1 ARPYSF 6

RESULT 3
ALL4_CALVO STANDARD; PRT; 8 AA.
AC P41840;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CALLATOSTATIN 4 (LEU-CALLATOSTATIN 4).
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
RN [1]
RP SEQUENCE
RC TISSUE=THORACIC GANGLION;
RX MEDLINE=932111980; PubMed=8460157;
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,
RA Thorpe A.;
RT "Callatostatin: neuropeptides from the blowfly Calliphora vomitoria
RT with sequence homology to cockroach allatostatins."
RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
RN [2]
RP CHARACTERIZATION.
RC TISSUE=THORACIC GANGLION;
RX MEDLINE=94291167; PubMed=8020069;
RA Duve H., Thorpe A.;
RT "Distribution and functional significance of Leu-callatostatins in
RT the blowfly Calliphora vomitoria."
RL Cell Tissue Res. 276:367-379(1994).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY
CC A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE
CC INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO
CC BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
CC FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.
CC -!- TISSUE SPECIFICITY: BRAIN, SUBESOPHAGEAL GANGLION, RETROCEREBRAL
CC COMPLEX, THORACICO-ABDOMINAL GANGLION, PERIPHERAL NEUROSECRETORY
CC SYSTEM AND INTESTINE.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
DR PIR; D47393; D47393.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 OR N.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 954 MW; D32879D5AB47740A CRC64;

Query Match 60.4%; Score 29; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPYSF 9
Db 2 RPYSF 6

RESULT 4
ALL0_CARMA STANDARD; PRT; 9 AA.
AC P61813;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

KW Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 963 MW; 372D79DCB4776C7 CRC64;

Query Match 52.1%; Score 25; DB 1; Length 9;
Best Local Similarity 57.1%; Pred. No. 8.9e+04;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 ASRPYSF 9
Db 1 AQPYPF 7

RESULT 5
ALL5_CARMA STANDARD; PRT; 8 AA.
AC P81818;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CARCINUSTATIN 15.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
RN [1]
RP SEQUENCE.
RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatatin superfamily in the shore crab Carcinus maenas."
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 963 MW; 372D79DCB4776C7 CRC64;

Query Match 50.0%; Score 24; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9
Db 3 PYSF 6

RESULT 6
ALL3_CARMA STANDARD; PRT; 7 AA.
AC P81806;
DT 30-MAY-2000 (Rel. 39, Created)

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE CARCINUSTATIN 3.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 796 MW; 672879DCB476B70 CRC64;

Query Match 43.8%; Score 21; DB 1; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9  
 II:I  
 Db 2 PYAF 5

RESULT 7  
 ALL4\_CARMA  
 ID ALL4\_CARMA STANDARD; PRT; 7 AA.  
 AC P81807;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 4.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 782 MW; 672879DCB476AC0 CRC64;

Query Match 43.8%; Score 21; DB 1; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9  
 II:I  
 Db 2 PYAF 5

RESULT 8  
 ALL5\_CARMA  
 ID ALL5\_CARMA STANDARD; PRT; 7 AA.  
 AC P81808;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 5.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 781 MW; 672879DCB476420 CRC64;

Query Match 43.8%; Score 21; DB 1; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9  
 II:I  
 Db 2 PYAF 5

RESULT 9  
 ALL6\_CARMA  
 ID ALL6\_CARMA STANDARD; PRT; 8 AA.  
 AC P81819;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 16.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 FT MOD\_RES 8  
 SQ SEQUENCE 8 AA; 813 MW; 7C286B45AB476878 CRC64;

Query Match 43.8%; Score 21; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9  
 II:I  
 Db 3 PYSY 6

RESULT 10  
 ALL7\_CARMA  
 ID ALL7\_CARMA STANDARD; PRT; 8 AA.  
 AC P81809; P81810; P81804;  
 DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 7 [CONTAINS: CARCINUSTATIN 6; CARCINUSTATIN 1].  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 CC Neuropeptide; Amidation; Multigene family.  
 KW CHAIN 1 8 CARCINUSTATIN 7.  
 FT CHAIN 2 8 CARCINUSTATIN 6.  
 FT CHAIN 4 8 CARCINUSTATIN 1.  
 FT MOD\_RES 8 8 AMIDATION.  
 FT SEQUENCE 8 AA; 825 MW; 922879CDCB4775BD CRC64;

Query Match 43.8%; Score 21; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9  
 |||  
 Db 3 PYAF 6

# RESULT 11

ALL8\_CARMA  
 ID ALL8\_CARMA STANDARD; PRT; 8 AA.  
 AC P81811;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 8.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 CC Neuropeptide; Amidation; Multigene family.  
 KW MOD\_RES 8 8 AMIDATION.  
 FT SEQUENCE 8 AA; 795 MW; 922879CDCB47687D CRC64;

Query Match 43.8%; Score 21; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9  
 |||  
 Db 3 PYAF 6

# RESULT 12

ALL9\_CARMA  
 ID ALL9\_CARMA STANDARD; PRT; 8 AA.  
 AC P81812;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 9.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 CC Neuropeptide; Amidation; Multigene family.  
 KW MOD\_RES 8 8 AMIDATION.  
 FT SEQUENCE 8 AA; 781 MW; 7C2879CDCB476878 CRC64;

Query Match 43.8%; Score 21; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9  
 |||  
 Db 3 PYAF 6

# RESULT 13

ALL5\_CALVO  
 ID ALL5\_CALVO STANDARD; PRT; 8 AA.  
 AC P41841;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CALLATOSTATIN 5 (MET-CALLATOSTATIN 1) ((HYP3)MET-CALLATOSTATIN).  
 OS Calliphora vomitoria (Blue blowfly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Calliphoridae; Calliphora.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93211980; PubMed=8460157;  
 RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,  
 RA Thorpe A.;  
 RT "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria  
 RT with sequence homology to cockroach allatostatins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).  
 RN [2]  
 RP CHARACTERIZATION, AND HYDROXYLATION.  
 RC TISSUE=HEAD;  
 RX MEDLINE=94342269; PubMed=8063725;  
 RA Duve H., Johnsen A.H., Scott A.G., East P., Thorpe A.;  
 RT "[Hyp3]Met-callatostatin. Identification and biological properties of  
 RT a novel neuropeptide from the blowfly Calliphora vomitoria.";  
 RL J. Biol. Chem. 269:21059-21066(1994).  
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY  
 CC A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE  
 CC INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO  
 CC BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO  
 CC FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.  
 CC -1- TISSUE SPECIFICITY: NEURONS WITHIN BRAIN AND ABDOMINAL GANGLION.  
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 DR PIR; E47393; E47393.  
 KW Neuropeptide; Amidation; Hydroxylation.



FT MOD\_RES 3 3 HYDROXYLATION (IN 20% OF THE PEPTIDES).  
FT MOD\_RES 8 8 AMIDATION.  
SQ SEQUENCE 8 AA; 883 MW; 7D9879CABB477768 CRC64;

Query Match 41.7%; Score 20; DB 1; Length 8;  
Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9  
| | |  
Db 3 PYDF 6

RESULT 14  
FLA2\_TREHY STANDARD; PRT; 9 AA.  
AC P80159;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE FLAGELLAR FILAMENT OUTER LAYER PROTEIN FLAA2 (35 KDA SHEATH PROTEIN)  
DE (FRAGMENT).  
GN FLAA2.  
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Brachyspira.  
RN [1]  
RP SEQUENCE.  
RC STRAIN=C5;  
RX MEDLINE=93139764; PubMed=1487733;  
RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,  
van der Zeijst B.A.M., Kusters J.G.;  
RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are  
composed of two sheath proteins and three core proteins.";  
RL J. Gen. Microbiol. 138:2697-2706(1992).  
CC -1- FUNCTION: COMPONENT OF THE OUTER LAYER OF THE FLAGELLA.  
CC -1- SUBUNIT: OUTER LAYER IS COMPOSED OF TWO SHEATH PROTEINS, FLAA1  
(44 KDA) AND FLAA2 (35 KDA) AND A CORE THAT CONTAINS THREE  
CC PROTEINS FLAB1 (37 KDA), FLAB2 (34 KDA) AND FLAB3 (32 KDA).  
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC FLAGELLUM.  
KW Flagella; Periplasmic.  
FT UNSURE 2 2  
FT UNSURE 8 9  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 1129 MW; 855A19C68B4772D1 CRC64;

Query Match 39.6%; Score 19; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9  
| | |  
Db 4 PYMF 7

RESULT 15  
AL17\_CARMA STANDARD; PRT; 8 AA.  
AC P81820;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CARCINUSTATIN 17.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,

RA Thorpe A.;  
RT "Isolation and identification of multiple neuropeptides of the  
allatostatin superfamily in the shore crab Carcinus maenas.";  
RL Eur. J. Biochem. 250:727-734(1997).  
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation; Multigene family.  
FT MOD\_RES 8 8 AMIDATION (POTENTIAL).  
SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 37.5%; Score 18; DB 1; Length 8;  
Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 SRPYSF 9  
| | | |  
Db 1 SGQYSF 6

Search completed: January 17, 2001, 13:47:09  
Job time: 278 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:40:21 ; Search time 31.16 Seconds  
(without alignments)  
5.187 Million cell updates/sec

Title: US-08-765-837-4  
Perfect score: 48  
Sequence: 1 NQASRPYSF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues  
Total number of hits satisfying chosen parameters: 44977

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/PctUS\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	29	60.4	8	1 US-08-522-326-10	Sequence 10, Appl
2	29	60.4	8	1 US-08-522-326-15	Sequence 15, Appl
3	29	60.4	8	2 US-08-714-053-7	Sequence 7, Appl
4	21	43.8	8	2 US-08-373-190-47	Sequence 47, Appl
5	21	43.8	8	2 US-08-438-190A-47	Sequence 47, Appl
6	21	43.8	8	3 US-08-350-215-47	Sequence 47, Appl
7	21	43.8	8	3 US-09-287-145A-47	Sequence 47, Appl
8	20	41.7	4	5 5494663-15	Patent No. 5494663
9	20	41.7	6	2 US-08-714-053-3	Sequence 3, Appl
10	20	41.7	7	2 US-08-194-981E-65	Sequence 65, Appl
11	20	41.7	8	1 US-08-522-326-11	Sequence 11, Appl
12	20	41.7	8	1 US-08-522-326-16	Sequence 16, Appl
13	20	41.7	8	2 US-08-714-053-1	Sequence 1, Appl
14	20	41.7	8	2 US-08-714-053-8	Sequence 8, Appl
15	20	41.7	9	1 US-08-190-788A-287	Sequence 287, App
16	20	41.7	9	1 US-08-383-474B-288	Sequence 288, App
17	20	41.7	9	1 US-08-465-391A-287	Sequence 287, App
18	20	41.7	9	2 US-08-464-538B-286	Sequence 286, App
19	19	39.6	5	1 US-08-022-381A-9	Sequence 9, Appl
20	19	39.6	5	1 US-08-475-827A-9	Sequence 9, Appl
21	19	39.6	5	5 5215888-6	Patent No. 5215888
22	19	39.6	6	1 US-07-620-410-1	Sequence 1, Appl
23	19	39.6	6	1 US-07-920-878-1	Sequence 1, Appl
24	19	39.6	6	1 US-07-920-878-6	Sequence 6, Appl
25	19	39.6	6	1 US-07-920-878-11	Sequence 11, Appl
26	19	39.6	6	1 US-07-920-878-16	Sequence 16, Appl
27	19	39.6	6	1 US-07-920-878-21	Sequence 21, Appl
28	19	39.6	6	1 US-07-920-878-26	Sequence 26, Appl

Sequence 31, Appl  
Sequence 32, Appl  
Sequence 33, Appl  
Sequence 34, Appl  
Sequence 13, Appl  
Sequence 72, Appl  
Sequence 110, Appl  
Sequence 111, Appl  
Sequence 1, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 5, Appl  
Sequence 6, Appl  
Sequence 9, Appl  
Sequence 16, Appl

19 39.6 6 1 US-07-920-878-31  
19 39.6 6 1 US-07-920-878-32  
19 39.6 6 1 US-07-920-878-33  
19 39.6 6 1 US-07-920-878-34  
19 39.6 6 1 US-08-082-847-13  
19 39.6 6 1 US-08-014-979-72  
19 39.6 6 1 US-08-014-979-110  
19 39.6 6 1 US-08-014-979-111  
19 39.6 6 1 US-08-218-608-1  
19 39.6 6 2 US-08-672-805-17  
19 39.6 6 2 US-08-672-805-18  
19 39.6 6 2 US-08-776-613-1  
19 39.6 6 2 US-08-776-613-2  
19 39.6 6 2 US-08-776-613-5  
19 39.6 6 2 US-08-776-613-6  
19 39.6 6 2 US-08-776-613-9  
19 39.6 6 4 PCT-US95-02131-16

## ALIGNMENTS

RESULT 1  
US-08-522-326-10  
; Sequence 10, Application US/08522326  
; Patent No. 5753615  
; GENERAL INFORMATION:  
; APPLICANT: THORPE, Alan  
; APPLICANT: DUVE, Hanne  
; APPLICANT: JOHNSON, Anders Holten  
; TITLE OF INVENTION: NEUROPEPTIDES AND THEIR USE AS  
; TITLE OF INVENTION: INSECTICIDES  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDING ADDRESS:  
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: US  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/522,326  
; FILING DATE: 12-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305120.9  
; FILING DATE: 12-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LUDWIG, STEVEN R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0623.0340000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2540  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Calliphora vomitoria  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 8  
; OTHER INFORMATION: /product= "OTHER"  
; OTHER INFORMATION: /note= "C-terminal amide"  
; PUBLICATION INFORMATION:

; DOCUMENT NUMBER: WO 94/20530 A2  
; FILING DATE: 14-MAR-1994  
; PUBLICATION DATE: 15-SEP-1994  
US-08-522-326-10

Query Match 60.4%; Score 29; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPYSF 9  
| | | | |  
Db 2 RPYSF 6

RESULT 2  
US-08-522-326-15  
; Sequence 15, Application US/08522326  
; Patent No. 5753615  
; GENERAL INFORMATION:  
; APPLICANT: THORPE, Alan  
; APPLICANT: DUVE, Hanne  
; APPLICANT: JOHNSON, Anders Holten  
; TITLE OF INVENTION: NEUROPEPTIDES AND THEIR USE AS  
; TITLE OF INVENTION: INSECTICIDES  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: US  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/522,326  
; FILING DATE: 12-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305120.9  
; FILING DATE: 12-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LUDWIG, STEVEN R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0623.0340000  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-522-326-15

Query Match 60.4%; Score 29; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPYSF 9  
| | | | |  
Db 2 RPYSF 6

RESULT 3  
US-08-714-053-7  
; Sequence 7, Application US/08714053

; Patent No. 5863763  
; GENERAL INFORMATION:  
; APPLICANT: Duve, Hanne  
; APPLICANT: Thorpe, Alan  
; APPLICANT: JOHNSON, Anders H.  
; TITLE OF INVENTION: Insect Neuropeptides  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Avenue, NW, Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/714,053  
; FILING DATE: 09-SEP-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/00505  
; FILING DATE: 09-MAR-1995  
; APPLICATION NUMBER: WO PCT/GB95/00505  
; FILING DATE: 09-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9404529.1  
; FILING DATE: 09-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 0623.0430000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Calliphora vomitoria  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 8  
; OTHER INFORMATION: /note= "Amidated leucine residue at  
; OTHER INFORMATION: position 8"  
US-08-714-053-7

Query Match 60.4%; Score 29; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RPYSF 9  
| | | | |  
Db 2 RPYSF 6

RESULT 4  
US-08-373-190-47  
; Sequence 47, Application US/08373190  
; Patent No. 5851829  
; GENERAL INFORMATION:  
; APPLICANT: MARASCO, WAYNE  
; APPLICANT: HASELTINE, WILLIAM  
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF PROTEINS  
; NUMBER OF SEQUENCES: 79

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,190
; FILING DATE: 17-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06735
; FILING DATE: 16-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 41956-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: STRE UR 2002
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; US-08-373-190-47

```

```

Query Match 43.8%; Score 21; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NOASRP 6
Db 2 NEASYP 7

```

```

RESULT 5
US-08-438-190A-47
; Sequence 47, Application US/08438190A
; Patent No. 5965371
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE
; APPLICANT: HASELTINE, WILLIAM
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,190A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: EISENSTEIN, RONALD I.
; REGISTRATION NUMBER: 30628
; REFERENCE/DOCKET NUMBER: 41956
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-438-190A-47

```

```

Query Match 43.8%; Score 21; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NOASRP 6
Db 2 NEASYP 7

```

```

RESULT 6
US-08-350-215-47
; Sequence 47, Application US/08350215
; Patent No. 6004940
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE A.
; APPLICANT: RICHARDSON, JENNIFER
; TITLE OF INVENTION: INTRACELLULAR TARGETING OF ENDOGENOUS
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/350,215
; FILING DATE: 12-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: EISENSTEIN, RONALD I.
; REGISTRATION NUMBER: 30628
; REFERENCE/DOCKET NUMBER: 41956-CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-350-215-47

```

```

Query Match 43.8%; Score 21; DB 3; Length 8;

```

```
;
; FILING DATE: 02-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 954,418
; FILING DATE: 30-SEP-1992
; APPLICATION NUMBER: 496,800
; FILING DATE: 21-MAR-1990
; APPLICATION NUMBER: 812,796
; FILING DATE: 23-DEC-1985
; SEQ ID NO:15:
; LENGTH: 4
5494663-15

Query Match 41.7%; Score 20; DB 5; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PYSF 9
Db 1 PFSF 4

RESULT 9
US-09-287-145A-47
; Sequence 47, Application US/09287145A
; Patent No. 6072036
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE
; APPLICANT: HASELTINE, WILLIAM
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,145A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,190
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: EISENSTEIN, RONALD I.
; REGISTRATION NUMBER: 30628
; REFERENCE/DOCKET NUMBER: 41956
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-287-145A-47

Query Match 43.8%; Score 21; DB 3; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NQASRP 6
Db 2 NEASYP 7

RESULT 8
5494663-15
; Patent No. 5494663
; APPLICANT: YAMADA, MASAOKI; FURUTANI, YASUJI; YAMAYOSHI, MICHIKO;
; NOTAKE, MITSUE; YAMAGISHI, JUNICHI
; TITLE OF INVENTION: TREATMENT OF MICROBIAL INFECTION WITH
; INTERLEUKIN 1 POLYPEPTIDES
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,826
```

NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note= "Amidated methionine residue  
OTHER INFORMATION: at position 6"  
US-08-714-053-3

Query Match 41.7%; Score 20; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PVSF 9  
DB 1 PVDF 4

RESULT 10  
US-08-194-981E-65  
Sequence 65, Application US/08194981E  
Patent No. 5886157  
GENERAL INFORMATION:  
APPLICANT: GUENGERICH, F. Peter  
APPLICANT: GUO, Zuyu  
APPLICANT: SANDHU, Punam  
APPLICANT: GILLAM, Elizabeth M. J.  
TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF  
TITLE OF INVENTION: HUMAN  
TITLE OF INVENTION: CYTOCHROME P450  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: Suite 1200, 127 Peachtree Street, NE  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303-1811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/194,981E  
FILING DATE: February 10, 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth Selby  
REGISTRATION NUMBER: 38,298  
REFERENCE/DOCKET NUMBER: 22000.0022  
TELEPHONE: (404) 688-0770  
TELEFAX: (404) 688-9880  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
US-08-194-981E-65

Query Match 41.7%; Score 20; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ASRP 6  
DB 2 ASRP 5

RESULT 11  
US-08-522-326-11  
Sequence 11, Application US/08522326  
Patent No. 5753615  
GENERAL INFORMATION:  
APPLICANT: THORPE, Alan  
APPLICANT: DUVE, Hanne  
APPLICANT: JOHNSON, Anders Holten  
TITLE OF INVENTION: NEUROPEPTIDES AND THEIR USE AS  
TITLE OF INVENTION: INSECTICIDES  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: US  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/522,326  
FILING DATE: 12-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9305120.9  
FILING DATE: 12-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: LUDWIG, STEVEN R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0623.0340000  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Calliphora vomitoria  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 8  
OTHER INFORMATION: /product= "Other"  
OTHER INFORMATION: /note= "C-terminal amide"  
PUBLICATION INFORMATION:  
DOCUMENT NUMBER: WO 94/20530 A2  
FILING DATE: 14-MAR-1994  
PUBLICATION DATE: 15-SEP-1994  
US-08-522-326-11

Query Match 41.7%; Score 20; DB 1; Length 8;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PVSF 9  
DB 3 PVDF 6

RESULT 12  
US-08-522-326-16  
Sequence 16, Application US/08522326  
Patent No. 5753615  
GENERAL INFORMATION:  
APPLICANT: THORPE, Alan  
APPLICANT: DUVE, Hanne

APPLICANT: JOHNSEN, Anders Holten  
TITLE OF INVENTION: NEUROPEPTIDES AND THEIR USE AS  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: US  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/522,326  
FILING DATE: 12-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9305120.9  
FILING DATE: 12-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: LUDWIG, STEVEN R.  
REGISTRATION NUMBER: 36,203  
REFERENCE/DOCKET NUMBER: 0623.0340000  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-522-326-16

Query Match 41.7%; Score 20; DB 1; Length 8;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9  
Db 3 PYDF 6

RESULT 13  
US-08-714-053-1  
Sequence 1, Application US/08714053  
Patent No. 5863763  
GENERAL INFORMATION:  
APPLICANT: Duve, Hanne  
APPLICANT: Thorpe, Alan  
APPLICANT: Johnsen, Anders H.  
TITLE OF INVENTION: Insect Neuropeptides  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, NW, Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/714,053  
FILING DATE: 09-SEP-1996

CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/00505  
FILING DATE: 09-MAR-1995  
APPLICATION NUMBER: WO PCT/GB95/00505  
FILING DATE: 09-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9404529.1  
FILING DATE: 09-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0623.0430000  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Calliphora vomitoria  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 2  
OTHER INFORMATION: /note= "Hydroxyproline residue at  
OTHER INFORMATION: position 2"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 8  
OTHER INFORMATION: /note= "Amidated methionine residue  
OTHER INFORMATION: at position 8"  
US-08-714-053-1

Query Match 41.7%; Score 20; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9  
Db 3 PYDF 6

RESULT 14  
US-08-714-053-8  
Sequence 8, Application US/03714053  
Patent No. 5863763  
GENERAL INFORMATION:  
APPLICANT: Duve, Hanne  
APPLICANT: Thorpe, Alan  
APPLICANT: Johnsen, Anders H.  
TITLE OF INVENTION: Insect Neuropeptides  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, NW, Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/714,053  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 530



;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/GB95/00505  
;; FILING DATE: 09-MAR-1995  
;; APPLICATION NUMBER: WO PCT/GB95/00505  
;; FILING DATE: 09-MAR-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: GB 9404529.1  
;; FILING DATE: 09-MAR-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Esmond, Robert W.  
;; REGISTRATION NUMBER: 32,893  
;; REFERENCE/DOCKET NUMBER: 0623.0430000  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-2600  
;; TELEFAX: 202-371-2540  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 8 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; ORIGINAL SOURCE:  
;; ORGANISM: Calliphora vomitoria  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 8  
;; OTHER INFORMATION: /note= "Amidated methionine residue  
;; OTHER INFORMATION: at position 8"  
;; US-08-714-053-8

Query Match 41.7%; Score 20; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PYSF 9  
  111  
DB 3 PYDF 6

RESULT 15  
US-08-190-788A-287  
; Sequence 287, Application US/08190788A  
; Patent No. 5608035  
; GENERAL INFORMATION:  
; APPLICANT: Yanofsky, Stephen D.  
; APPLICANT: Barrett, Ronald W.  
; APPLICANT: Baldwin, David N.  
; APPLICANT: Jacobs, Jeff W.  
; TITLE OF INVENTION: Peptides and Compounds That Bind to the  
; TITLE OF INVENTION: IL-1 Receptor  
; NUMBER OF SEQUENCES: 312  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Affymax Technologies N.V.  
; STREET: 4001 Miranda Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,788A  
; FILING DATE: 02-FEB-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847,567  
; FILING DATE: 05-MAR-1992  
; ATTORNEY/AGENT INFORMATION:

;; NAME: Stevens, Lauren L.  
;; REGISTRATION NUMBER: 36,691  
;; REFERENCE/DOCKET NUMBER: 1019.1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-496-2300  
;; TELEFAX: 415-424-0832  
;; INFORMATION FOR SEQ ID NO: 287:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Region  
;; LOCATION: 3  
;; OTHER INFORMATION: /note= "naphthylalanine."  
;; US-08-190-788A-287

Query Match 41.7%; Score 20; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRPYS 8  
  :||||  
DB 3 AQPYS 7

Search completed: January 17, 2001, 13:42:56  
Job time: 155 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:43:01 ; Search time 14.14 Seconds  
(without alignments)  
16.928 Million cell updates/sec

Title: US-08-765-837-7  
Perfect score: 43  
Sequence: 1 ETKSWYF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 35323

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_36:\*\*  
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	7	R90951	Factor VIII antige
2	24	55.8	4	R38120	Protease-non-labil
3	24	55.8	4	W43253	Peptide #30 having
4	24	55.8	6	W28886	Oploid peptide. S
5	24	55.8	6	W15206	Anti-adherence pep
6	24	55.8	6	W15207	Anti-adherence pep
7	24	55.8	6	W15208	Anti-adherence pep
8	24	55.8	6	W15209	Anti-adherence pep
9	24	55.8	6	W15210	Anti-adherence pep
10	24	55.8	6	W15211	Anti-adherence pep
11	24	55.8	6	W15212	Anti-adherence pep
12	24	55.8	6	R93744	New peptide which

13	24	55.8	6	19	W40931	Leader sequence 50
14	24	55.8	6	20	Y22992	Oploid peptide whi
15	24	55.8	7	19	W40932	Leader sequence 51
16	24	55.8	7	19	W40933	Leader sequence 12
17	24	55.8	7	20	W40879	Puumala virus Inhi
18	24	55.8	7	20	Y49051	Membrane dipeptida
19	23	53.5	4	17	W04819	Immunomodulatory p
20	23	53.5	6	18	W28978	Oploid peptide. S
21	23	53.5	6	18	W24339	New peptide which
22	23	53.5	6	20	Y23089	Oploid peptide whi
23	23	53.5	6	21	Y83833	Pyroglutaminolazep
24	23	53.5	7	20	Y30217	Physarum polyceph
25	22	51.2	4	19	W63815	Human IL-18 recept
26	22	51.2	4	20	Y39794	Interleukin-18 rec
27	22	51.2	5	12	R12731	Pentapeptide paral
28	22	51.2	5	16	R69963	Pentameric mimotop
29	22	51.2	5	17	R98709	Control peptide fo
30	22	51.2	6	16	R79636	Endocarditis spec
31	22	51.2	6	20	Y33848	G30 sub domain of
32	21	48.8	5	18	W22916	Low density lipopr
33	21	48.8	5	20	W78366	Isoelectric focus
34	21	48.8	6	18	W28990	Oploid peptide. S
35	21	48.8	6	18	W23133	Cyclic somatostati
36	21	48.8	6	18	R93748	New peptide which
37	21	48.8	6	18	W03277	Viral integrase in
38	21	48.8	6	19	W66594	Peptide component
39	21	48.8	6	20	Y25854	Human secreted pro
40	21	48.8	6	20	Y22996	Oploid peptide whi
41	21	48.8	6	20	W78367	Isoelectric focus
42	21	48.8	6	21	Y91881	Internal peptide o
43	21	48.8	6	21	Y91882	Internal peptide o
44	21	48.8	7	15	R62543	Fibronectin gelati
45	21	48.8	7	20	Y42870	Vpr-binding peptid

## ALIGNMENTS

RESULT 1	
R90951	ID R90951 standard; peptide; 7 AA.
XX	AC R90951;
XX	DT 09-SEP-1996 (first entry)
XX	DE Factor VIII antigenic peptide corresp. to residues Glu1885-Phe1891.
XX	KW Factor VIII; modification; inhibitor activity; binding: antibody;
XX	KW von Willebrand factor; immune disorder.
XX	OS Synthetic.
XX	PN W09602572-A2.
XX	PD 01-FEB-1996.
XX	PF 14-JUL-1995; 95WO-BE000068.
XX	PR 14-JUL-1994; 94BE-0000666.
XX	(CROI-) CROIX ROUGE BELGIOUE.
XX	PI Di Giambattista M, Laub R;
XX	DR Anti-adherence pep
XX	XX Anti-adherence pep
XX	PT Anti-adherence pep
XX	PT Anti-adherence pep
XX	PT Anti-adherence pep
XX	PT Anti-adherence pep
XX	XX Anti-adherence pep
XX	PS Claim 8; Page 31; 45pp; French.

```

XX Peptides R0945-64 are derived from the factor VIII protein, esp. from a
CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,
CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII
CC and derived peptides can be used to prevent the activity of inhibitors of
CC factor VIII binding to von Willebrand factor, esp. antibodies, thus
CC preventing or treating immune disorders.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 17; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETKSWYF 7
DB 1 etkswyf 7
   |||||

RESULT 2
R38120
ID R38120 standard; peptide; 4 AA.
XX
AC R38120;
XX
XX 12-OCT-1993 (first entry)
XX
XX Protease-non-labile fragment #17.
XX
XX Protease-labile; lipase; Humicola lanuginosa; strain DSM 4106;
KW non-labile; stability; proteolytic degradation; detergent; protease.
XX
XX Synthetic.
XX
XX W09311254-A.
XX
XX 10-JUN-1993.
XX
XX 26-NOV-1992; 92WO-DK00351.
XX
XX 26-NOV-1991; 91WO-DK00350.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Clausen IG, Pathar SA, Svendsen A;
XX
XX WPI; 1993-197071/24.
XX
XX Protein with improved stability against proteolytic degradation -
PT comprises one or more protease labile aminoacid segments, substd.
PT by non-labile segments used in detergents in conjunction with a
PT protease
XX
XX Claim 10; Page 44; 55pp; English.
XX
XX The sequences given in R38104-21 represent protease-non-labile amino
CC acid segments which may be substituted into the wild-type lipase
CC protein of Humicola lanuginosa strain DSM 4106 at position 209-212
CC (see also R38101). The lipase resulting from these substitutions has
CC improved stability against proteolytic degradation and is suitable
CC for use in detergent compositions especially when used in conjunction
CC with a protease.
XX
XX Sequence 4 AA;

Query Match 55.8%; Score 24; DB 14; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7
   |||
DB 2 wyf 4

RESULT 4
W28886
ID W28886 standard; peptide; 6 AA.
XX

```

```

Db 1 wyf 3

RESULT 3
W43253
ID W43253 standard; peptide; 4 AA.
XX
AC W43253;
XX
XX 31-MAR-1998 (first entry)
XX
XX Peptide #30 having anti-microbial activity.
XX
XX Anti-microbial peptide; bee sting; melittin; haemolytic activity;
KW bee venom; red blood cell; cell lysis; inhibition; Escherichia Coli;
KW trypsin; D-amino acid; IC50.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "fluorenylmethoxycarbonyl-modified lysine"
FT Modified-site 4
FT /note= "D-form residue, C-terminal amide"
XX
XX US5698673-A.
XX
XX 16-DEC-1997.
XX
XX 04-MAY-1995; 95US-0434761.
XX
XX 18-JUN-1993; 93US-0079445.
XX
XX (TORR-) TORREY PINES INST.
XX
XX Blondelle SE, Houghten RA, Pinilla C;
XX
XX WPI; 1998-051545/05.
XX
XX Anti-melittin peptide(s) - useful for treating bee stings
XX
XX Example 1; Column 18; 26pp; English.
XX
XX Peptides represented by the formula (Kfmoc)WYX, where X is any L-Or
CC D-amino acid and Kfmoc is fluorenylmethoxycarbonyl-lysine, were
CC synthesised and screened for anti-microbial activity against Escherichia
CC coli. The present peptide showed anti-microbial activity with an IC50
CC value (concentration necessary to inhibit 50 per cent growth of bacteria)
CC of 116 microg/ml. The invention provides peptides having anti-melittin
CC activity with IC50 values of less than 30 microg/ml which are useful for
CC treating melittin poisoning. Melittin is the active compound in bee venom
CC and causes cell lysis. Peptides having anti-melittin activity have one of
CC the following formulae; AC-IVILH22-NH2; AC-IVILH22-NH2; AC-IVILH22-NH2;
CC AC-IVIFEZ-NH2; AC-Z1-Z2-Z3-Z4-E-NH2; AC-MILWIE-NH2; AC-VIQQEV-NH2;
CC AC-WIQIFR-NH2; where Z= an optionally methylated or chlorinated D-form or
CC natural amino acid; Z1= F or I; Z2= W or Q; Z3= W or Y and Z4= C or F.
XX
XX Sequence 4 AA;

Query Match 55.8%; Score 24; DB 19; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7
   |||
DB 2 wyf 4

RESULT 4
W28886
ID W28886 standard; peptide; 6 AA.
XX

```

AC W28886;  
 XX 20-JAN-1998 (first entry)  
 DT  
 XX Opioid peptide.  
 DE  
 XX enkephalin; mu-opioid receptor ligand; agonist; antagonist.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-acetyl-Phe"  
 FT Modified-site 6  
 FT /note= "the C-terminal is in amide form"  
 FT  
 XX US5641861-A.  
 PN  
 XX 24-JUN-1997.  
 PD  
 XX 07-JUN-1995; 95US-0487006.  
 PF  
 XX 07-JUN-1995; 95US-0487006.  
 PR  
 XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.  
 PA  
 XX Dooley CT, Houghten RA;  
 PI  
 XX WPI; 1997-340994/31.  
 DR  
 XX New opioid peptides which bind mu receptors specifically - have  
 XX agonist or antagonist activity and are used for study and  
 PT localisation of mu receptors and to treat peripheral side effects of  
 PT morphine etc.  
 PT  
 XX Disclosure; Column 8; 92pp; English.  
 PS  
 XX The patent discloses the following new peptides, which are opioids which  
 XX bind specifically to the mu receptor: Ac-Phe-Arg-Trp-Trp-Tyr-Xaa-NH2 (1);  
 CC Ac-Arg-Trp-Ile-Gly-Trp-Xaa-NH2 (2); Trp-Trp-Pro-Lys-His-Xaa-NH2 (3);  
 CC Trp-Trp-Pro-Xaa-NH2 (4); Tyr-Pro-Phe-Gly-Xaa-NH2 (5);  
 CC D-Ile-D-Met-D-Ser-D-Trp-D-Trp-(Gly)n-Xaa2-NH2 (6);  
 CC D-Ile-D-Met-D-Thr-D-Trp-Gly-Xaa2-NH2 (7); Tyr-Al-B2-C3-NH2 (214);  
 CC Pm and red ((Me)x(H)y-Tyr-(NMe)z-Tyr-(Xaa3)z-NH2) (221); and  
 CC Trp-Trp-Pro-D4-(His)z-(Xaa)z-NH2 (222); where Xaa = any natural amino  
 CC acid; Xaa1 = Lys or Arg; n and z = 0 or 1; Xaa2 = Gly or the D form of  
 CC any naturally occurring amino acid; A1 = D-norvaline or D-norleucine;  
 CC B2 = Gly, Phe or Trp; C3 = Trp or naphthylalanine; x and y = 0-2, but  
 CC not over 2 in total; Xaa3 = Phe, Dphe or benzylamino; D4 = Lys or Arg;  
 CC Pm and red indicate permethylation and reduction of all CO in peptide  
 CC links to methylene. These new compounds are useful: (i) for in vitro  
 CC assay and study of opiate receptor subtypes, particularly mu receptors  
 CC in the brain; (ii) for in vivo localisation of receptor subtypes; and  
 CC (iii) therapeutically to block the peripheral effects (e.g. constipation  
 CC and pruritus) of centrally acting pain killers such as morphine.  
 CC They are very selective for the mu opioid receptor, over binding to the  
 CC delta and kappa receptor subtypes.  
 CC The present sequence is a specific example of peptide (1).  
 XX  
 XX Sequence 6 AA;  
 SQ  
 Query Match 55.8%; Score 24; DB 18; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 WYF 7  
 Db 4 WYF 6  
 RESULT 5  
 W15206

ID W15206 standard; peptide; 6 AA.  
 XX  
 AC W15206;  
 XX  
 DT 12-JAN-1998 (first entry)  
 XX  
 DE Anti-adherence peptide for use in oral care.  
 XX  
 KW Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;  
 KW periodontal disease; dental implant infection; dental caries;  
 KW bacteria; teeth.  
 OS Synthetic.  
 XX  
 PN GB2307476-A.  
 XX  
 PD 28-MAY-1997.  
 XX  
 XX 03-JUN-1996; 96GB-0011500.  
 PF  
 XX 06-JUN-1995; 95US-0466542.  
 PR  
 XX (PROC ) PROCTER & GAMBLE CO.  
 PA  
 XX Baker TR, Charbonneau DL, Murawski SL, Ward SR;  
 PI WPI; 1997-261853/24.  
 XX  
 DR New peptide(s) and their dimers, cyclo-monomer(s) and polymers -  
 XX useful for treating periodontal disease, plaque, dental implant  
 PT infection and dental caries  
 PT  
 XX Claim 1; Page 31; 42pp; English.  
 PS  
 XX The present sequence represents a peptide with anti-adherence activity.  
 CC The peptide can be used in an oral care composition, comprising the  
 CC peptide and an excipient. The peptide can be used to prevent or reduce  
 CC bacteria adherence to teeth and/or coaggregation between plaque forming  
 CC bacteria. It is useful for reducing or preventing periodontal disease,  
 CC plaque, dental implant infection and dental caries when applied  
 CC topically to tooth surfaces and/or tissues of the oral cavity. An  
 CC effective dosage of the peptide is 0.02-0.1 g. Dimeric and  
 CC cyclomonomeric anti-adherence peptides exhibit enhanced stability to  
 CC the acidic conditions and proteases which exist in the oral cavity.  
 XX  
 XX Sequence 6 AA;  
 SQ  
 Query Match 55.8%; Score 24; DB 18; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 WYF 7  
 Db 1 WYF 3  
 RESULT 6  
 W15207  
 ID W15207 standard; peptide; 6 AA.  
 XX  
 AC W15207;  
 XX  
 DT 12-JAN-1998 (first entry)  
 XX  
 DE Anti-adherence peptide for use in oral care.  
 XX  
 KW Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;  
 KW periodontal disease; dental implant infection; dental caries;  
 KW bacteria; teeth.  
 OS Synthetic.  
 XX

PN GB2307476-A.  
 XX 28-MAY-1997.  
 XX 03-JUN-1996; 96GB-0011500.  
 XX 06-JUN-1995; 95US-0466542.  
 XX (PROC ) PROCTER & GAMBLE CO.  
 XX Baker TR, Charbonneau DL, Murawski SL, Ward SR;  
 XX WPI; 1997-261853/24.  
 XX New peptide(s) and their dimers, cyclo-monomer(s) and polymers -  
 PT useful for treating periodontal disease, plaque, dental implant  
 PT infection and dental caries  
 XX Claim 1; Page 31; 42pp; English.  
 XX The present sequence represents a peptide with anti-adherence activity.  
 CC The peptide can be used in an oral care composition, comprising the  
 CC peptide and an excipient. The peptide can be used to prevent or reduce  
 CC bacteria adherence to teeth and/or coaggregation between plaque forming  
 CC bacteria. It is useful for reducing or preventing periodontal disease,  
 CC plaque, dental implant infection and dental caries when applied  
 CC topically to tooth surfaces and/or tissues of the oral cavity. An  
 CC effective dosage of the peptide is 0.02-0.1 g. Dimeric and  
 CC cyclononameric anti-adherence peptides exhibit enhanced stability to  
 CC the acidic conditions and proteases which exist in the oral cavity.  
 XX SQ Sequence 6 AA;  
 XX Query Match 55.8%; Score 24; DB 18; Length 6;  
 XX Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 XX Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 WYF 7  
 Db 1 wyf 3  
 RESULT 7  
 W15208  
 ID W15208 standard; peptide; 6 AA.  
 AC W15208;  
 XX 12-JAN-1998 (first entry)  
 DT Anti-adherence peptide for use in oral care.  
 DE Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;  
 KW periodontal disease; dental implant infection; dental caries;  
 KW bacteria; teeth.  
 XX Synthetic.  
 OS GB2307476-A.  
 PN 28-MAY-1997.  
 XX 03-JUN-1996; 96GB-0011500.  
 XX 06-JUN-1995; 95US-0466542.  
 XX (PROC ) PROCTER & GAMBLE CO.  
 XX Baker TR, Charbonneau DL, Murawski SL, Ward SR;  
 XX WPI; 1997-261853/24.  
 XX New peptide(s) and their dimers, cyclo-monomer(s) and polymers -  
 PT useful for treating periodontal disease, plaque, dental implant  
 PT infection and dental caries  
 XX Claim 1; Page 31; 42pp; English.  
 XX The present sequence represents a peptide with anti-adherence activity.  
 CC The peptide can be used in an oral care composition, comprising the  
 CC peptide and an excipient. The peptide can be used to prevent or reduce  
 CC bacteria adherence to teeth and/or coaggregation between plaque forming  
 CC bacteria. It is useful for reducing or preventing periodontal disease,  
 CC plaque, dental implant infection and dental caries when applied  
 CC topically to tooth surfaces and/or tissues of the oral cavity. An  
 CC effective dosage of the peptide is 0.02-0.1 g. Dimeric and  
 CC cyclononameric anti-adherence peptides exhibit enhanced stability to  
 CC the acidic conditions and proteases which exist in the oral cavity.  
 XX SQ Sequence 6 AA;  
 XX Query Match 55.8%; Score 24; DB 18; Length 6;  
 XX Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 XX Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 WYF 7  
 Db 1 wyf 3  
 RESULT 7  
 W15208  
 ID W15208 standard; peptide; 6 AA.  
 AC W15208;  
 XX 12-JAN-1998 (first entry)  
 DT Anti-adherence peptide for use in oral care.  
 DE Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;  
 KW periodontal disease; dental implant infection; dental caries;  
 KW bacteria; teeth.  
 XX Synthetic.  
 OS GB2307476-A.  
 PN 28-MAY-1997.  
 XX 03-JUN-1996; 96GB-0011500.  
 XX 06-JUN-1995; 95US-0466542.  
 XX (PROC ) PROCTER & GAMBLE CO.  
 XX Baker TR, Charbonneau DL, Murawski SL, Ward SR;  
 XX WPI; 1997-261853/24.

PT New peptide(s) and their dimers, cyclo-monomer(s) and polymers -  
 PT useful for treating periodontal disease, plaque, dental implant  
 PT infection and dental caries  
 XX Claim 1; Page 32; 42pp; English.  
 XX The present sequence represents a peptide with anti-adherence activity.  
 CC The peptide can be used in an oral care composition, comprising the  
 CC peptide and an excipient. The peptide can be used to prevent or reduce  
 CC bacteria adherence to teeth and/or coaggregation between plaque forming  
 CC bacteria. It is useful for reducing or preventing periodontal disease,  
 CC plaque, dental implant infection and dental caries when applied  
 CC topically to tooth surfaces and/or tissues of the oral cavity. An  
 CC effective dosage of the peptide is 0.02-0.1 g. Dimeric and  
 CC cyclononameric anti-adherence peptides exhibit enhanced stability to  
 CC the acidic conditions and proteases which exist in the oral cavity.  
 XX SQ Sequence 6 AA;  
 XX Query Match 55.8%; Score 24; DB 18; Length 6;  
 XX Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 XX Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 WYF 7  
 Db 1 wyf 3  
 RESULT 8  
 W15209  
 ID W15209 standard; peptide; 6 AA.  
 AC W15209;  
 XX 12-JAN-1998 (first entry)  
 DT Anti-adherence peptide for use in oral care.  
 DE Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;  
 KW periodontal disease; dental implant infection; dental caries;  
 KW bacteria; teeth.  
 XX Synthetic.  
 OS GB2307476-A.  
 PN 28-MAY-1997.  
 XX 03-JUN-1996; 96GB-0011500.  
 XX 06-JUN-1995; 95US-0466542.  
 XX (PROC ) PROCTER & GAMBLE CO.  
 XX Baker TR, Charbonneau DL, Murawski SL, Ward SR;  
 XX WPI; 1997-261853/24.  
 XX New peptide(s) and their dimers, cyclo-monomer(s) and polymers -  
 PT useful for treating periodontal disease, plaque, dental implant  
 PT infection and dental caries  
 XX Claim 1; Page 32; 42pp; English.  
 XX The present sequence represents a peptide with anti-adherence activity.  
 CC The peptide can be used in an oral care composition, comprising the  
 CC peptide and an excipient. The peptide can be used to prevent or reduce  
 CC bacteria adherence to teeth and/or coaggregation between plaque forming  
 CC bacteria. It is useful for reducing or preventing periodontal disease,  
 CC plaque, dental implant infection and dental caries when applied  
 CC topically to tooth surfaces and/or tissues of the oral cavity. An  
 CC effective dosage of the peptide is 0.02-0.1 g. Dimeric and

CC cyclomonomeric anti-adherence peptides exhibit enhanced stability to  
 XX the acidic conditions and proteases which exist in the oral cavity.

SQ Sequence 6 AA;

Query Match 55.8%; Score 24; DB 18; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.1e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7

|||

Db 1 wvf 3

RESULT 9

W15210

ID W15210 standard; peptide; 6 AA.

XX AC

XX W15210;

DT 12-JAN-1998 (first entry)

XX DE

XX Anti-adherence peptide for use in oral care.

XX Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;  
 KW periodontal disease; dental implant infection; dental caries;  
 KW bacteria; teeth.

OS Synthetic.

XX GB2307476-A.

PN 28-MAY-1997.

XX PF

XX 03-JUN-1996; 96GB-0011500.

XX PR

XX 06-JUN-1995; 95US-0466542.

XX PA

XX (PROC ) PROCTER & GAMBLE CO.

XX PI

XX Baker TR, Charbonneau DL, Murawski SL, Ward SR;

XX WPI; 1997-261853/24.

XX DR

XX New peptide(s) and their dimers, cyclo-monomer(s) and polymers -

XX useful for treating periodontal disease, plaque, dental implant

XX infection and dental caries

XX Claim 1; Page 32; 42pp; English.

XX The present sequence represents a peptide with anti-adherence activity.

XX The peptide can be used in an oral care composition, comprising the

XX peptide and an excipient. The peptide can be used to prevent or reduce

XX bacteria adherence to teeth and/or coaggregation between plaque forming

XX bacteria. It is useful for reducing or preventing periodontal disease,

XX plaque, dental implant infection and dental caries when applied

XX topically to tooth surfaces and/or tissues of the oral cavity. An

XX effective dosage of the peptide is 0.02-0.1 g. Dimeric and

XX cyclomonomeric anti-adherence peptides exhibit enhanced stability to

XX the acidic conditions and proteases which exist in the oral cavity.

XX SQ Sequence 6 AA;

Query Match 55.8%; Score 24; DB 18; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.1e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7

|||

Db 1 wvf 3

RESULT 11

W15212

ID W15212 standard; peptide; 6 AA.

XX AC

XX W15212;

DT 12-JAN-1998 (first entry)

XX DE

XX Anti-adherence peptide for use in oral care.

XX Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;

XX KW periodontal disease; dental implant infection; dental caries;

XX KW bacteria; teeth.

RESULT 10

W15211

ID W15211 standard; peptide; 6 AA.

XX AC

XX W15211;

DT 12-JAN-1998 (first entry)

XX DE

XX Anti-adherence peptide for use in oral care.

XX Anti adherence; oral care; dimer; cyclo monomer; polymer; plaque;

XX KW periodontal disease; dental implant infection; dental caries;

XX KW bacteria; teeth.

XX OS Synthetic.

XX GB2307476-A.

PN 28-MAY-1997.

XX PF

XX 03-JUN-1996; 96GB-0011500.

XX PR

XX 06-JUN-1995; 95US-0466542.

XX PA

XX (PROC ) PROCTER & GAMBLE CO.

XX PI

XX Baker TR, Charbonneau DL, Murawski SL, Ward SR;

XX WPI; 1997-261853/24.

XX DR

XX New peptide(s) and their dimers, cyclo-monomer(s) and polymers -

XX useful for treating periodontal disease, plaque, dental implant

XX infection and dental caries

XX Claim 1; Page 33; 42pp; English.

XX The present sequence represents a peptide with anti-adherence activity.

XX The peptide can be used in an oral care composition, comprising the

XX peptide and an excipient. The peptide can be used to prevent or reduce

XX bacteria adherence to teeth and/or coaggregation between plaque forming

XX bacteria. It is useful for reducing or preventing periodontal disease,

XX plaque, dental implant infection and dental caries when applied

XX topically to tooth surfaces and/or tissues of the oral cavity. An

XX effective dosage of the peptide is 0.02-0.1 g. Dimeric and

XX cyclomonomeric anti-adherence peptides exhibit enhanced stability to

XX the acidic conditions and proteases which exist in the oral cavity.

XX SQ Sequence 6 AA;

Query Match 55.8%; Score 24; DB 18; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.1e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7

|||

Db 1 wvf 3

XX OS Synthetic.  
 XX PN GB2307476-A.  
 XX PD 28-MAY-1997.  
 XX PF 03-JUN-1996; 96GB-0011500.  
 XX PR 06-JUN-1995; 95US-0466542.  
 XX PA (PROC ) PROCTER & GAMBLE CO.  
 XX PI Baker TR, Charbonneau DL, Murawski SL, Ward SR;  
 XX DR WPI; 1997-261853/24.  
 XX PT New peptide(s) and their dimers, cyclo-monomer(s) and polymers -  
 PT useful for treating periodontal disease, plaque, dental implant  
 PT infection and dental caries  
 XX PS Claim 1; Page 33; 42pp; English.  
 XX CC The present sequence represents a peptide with anti-adherence activity.  
 CC The peptide can be used in an oral care composition, comprising the  
 CC peptide and an excipient. The peptide can be used to prevent or reduce  
 CC bacteria adherence to teeth and/or coaggregation between plaque forming  
 CC bacteria. It is useful for reducing or preventing periodontal disease,  
 CC plaque, dental implant infection and dental caries when applied  
 CC topically to tooth surfaces and/or tissues of the oral cavity. An  
 CC effective dosage of the peptide is 0.02-0.1 g. Dimeric and  
 CC cyclomonomeric anti-adherence peptides exhibit enhanced stability to  
 CC the acidic conditions and proteases which exist in the oral cavity.  
 XX SQ Sequence 6 AA;

Query Match 55.8%; Score 24; DB 18; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7  
 DB 1 wyl 3

RESULT 12  
 R93744  
 ID R93744 standard; Protein; 6 AA.  
 AC R93744;  
 XX  
 DT 23-SEP-1997 (first entry)  
 XX  
 DE New peptide which acts as mu-opioid receptor ligand.  
 KW mu-receptor; opioid; opiate; agonist; antagonist; diagnosis;  
 KW analgesic.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-acetyl-Phe"  
 FT Misc-difference 6  
 FT /note= "this residue is in C-terminal amide form"  
 XX WO9640208-A1.  
 XX PN  
 XX PD 19-DEC-1996.  
 XX PF 06-JUN-1996; 96WO-US09321.  
 XX

PR 07-JUN-1995; 95US-0476438.  
 XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.  
 XX PI Dooley CT, Houghten RA;  
 XX DR WPI; 1997-051895/05.  
 XX PT New mu opioid receptor binding ligand peptide(s) - useful for  
 PT in-vitro and in-vivo diagnosis, as analgesics, and for blocking  
 PT peripheral effects of centrally acting drugs, e.g. morphine  
 XX PS Disclosure; Page 17; 57pp; English.  
 XX CC The patent discloses eight new groups of opioid peptides which bind  
 CC to the mu-receptor to act as agonists or antagonists. The peptides  
 CC can be used for in-vitro assays to study opiate receptor subtypes  
 CC (especially the mu type) in brain or other tissue samples; and for  
 CC in-vivo diagnosis to localise opioid subtypes. The peptides are also  
 CC useful as drugs to treat pathologies associated with other compounds  
 CC which interact with the opioid receptor system. Therefore they can be  
 CC used in medicaments for treating pathologies associated with the mu  
 CC receptor and as analgesics. They can be used therapeutically to block  
 CC the peripheral effects of centrally acting pain killers, e.g. lo  
 CC prevent side effects such as constipation and pruritis associated  
 CC with morphine. The present sequence represents a specific example  
 CC of one of the new groups of peptides, of formula  
 CC Ac-Phe-Arg-Trp-Trp-Tyr-Xaa-NH2 where Xaa = a naturally occurring  
 CC amino acid.  
 XX SQ Sequence 6 AA;

Query Match 55.8%; Score 24; DB 18; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7  
 DB 4 wyl 6

RESULT 13  
 W40931  
 ID W40931 standard; peptide; 6 AA.  
 XX AC W40931;  
 XX  
 DT 09-JUN-1998 (first entry)  
 XX  
 DE Leader sequence 50, development of substances against Bundyviridae.  
 XX  
 KW Virus inhibitory peptide; inhibition; ligand library;  
 KW competitive affinity-elution; prophylactic substance;  
 KW combinational therapy; pathogen diagnosis; leader molecule.  
 XX Bundyviridae.  
 OS  
 XX WO9745743-A1.  
 XX PN  
 XX PD 04-DEC-1997.  
 XX PF 30-MAY-1997; 97WO-FI00339.  
 XX PR 30-MAY-1996; 96FI-0002269.  
 XX (UYHE-) UNIV HELSINKI LICENSING LTD.  
 XX PA  
 XX Heiskanen T, Lankinen H, Lundkvist A, Vaheri A;  
 XX WPI; 1998-032807/03.  
 XX PT Selection of target pathogen inhibiting substances - useful for



PT comparative drug design to provide therapeutically active,  
 PT protective and prophylactic substances  
 PS Claim 16; Page 69; 81pp; English.  
 XX This is a peptide which is useful as a leader molecule. It is used  
 CC for the development of substances active against enveloped pathogens  
 CC such as viruses from the family Bunyaviridae. Its selection  
 CC involved the reaction of a ligand library with the bound target  
 CC pathogen and the pathogen bound ligands are subjected to competitive  
 CC affinity-elution with at least one neutralising substance. The method  
 CC is useful for comparative drug design to provide therapeutically active,  
 CC protective and/or prophylactic substances and developing combinational  
 CC therapies as well as for pathogen diagnostics. The method also  
 CC identifies "leader molecules", which not only have great affinity,  
 CC but also mimic the functions of known neutralising substances.  
 XX Sequence 6 AA;  
 SQ

Query Match 55.8%; Score 24; DB 19; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7  
 DB 1 wyf 3  
 III

RESULT 14  
 ID Y22992 standard; peptide; 6 AA.  
 AC Y22992;  
 XX 23-AUG-1999 (first entry)  
 DT  
 DE Opioid peptide which inhibits binding of enkephalin.  
 XX  
 KW Opioid peptide; ligand binding; opioid receptor;  
 KW micro-selective opioid peptide; enkephalin; opioid receptor system;  
 KW blocking; peripheral effect; centrally acting pain killer; morphine.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "acetylated"  
 FT Modified-site 6 /note= "amidated"  
 FT  
 XX US5919897-A.  
 XX  
 PD 06-JUL-1999.  
 XX  
 PF 07-JUN-1995; 95US-0488659.  
 XX  
 PR 07-JUN-1995; 95US-0488659.  
 XX  
 XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.  
 XX  
 PI Dooley CT, Houghten RA;  
 XX  
 DR WPI; 1999-394647/33.  
 XX  
 PT New opioid peptides useful for blocking the peripheral effects of  
 PT centrally acting pain killers such as morphine  
 XX  
 PS Example 1; Column 8; 92pp; English.  
 XX  
 CC The specification describes opioid peptides, in which each of the  
 CC N atoms in the peptide backbone between respective amino acids is  
 CC modified by permethylation, perallylation, perethylation, perbenzoylation

CC and pernapthylation. The peptides inhibit ligand binding to an opioid  
 CC receptor. Specifically, the peptides inhibit the micro-selective  
 CC opioid peptide enkephalin. The peptides can be used in vivo  
 CC diagnostically to localize opioid receptor subtypes. They can be used  
 CC to treat pathologies associated with other compounds which interact with  
 CC the opioid receptor system. The peptides are especially useful for  
 CC blocking the peripheral effects of centrally acting pain killers such  
 CC as morphine. Y22984-Y23003 represent opioid peptides of the invention,  
 CC and are derived from the general sequence given in Y22983.  
 XX Sequence 6 AA;  
 SQ

Query Match 55.8%; Score 24; DB 20; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7  
 DB 4 wyf 6  
 III

RESULT 15  
 ID W40932 standard; peptide; 7 AA.  
 AC W40932;  
 XX  
 DT 09-JUN-1998 (first entry)  
 DE  
 DE Leader sequence 51, development of substances against Bunyaviridae;  
 XX  
 KW Virus inhibitory peptide; inhibition; ligand library;  
 KW competitive affinity-elution; prophylactic substance;  
 KW combinational therapy; pathogen diagnosis; leader molecule.  
 XX  
 OS Bunyaviridae.  
 XX  
 PN WO9745743-A1.  
 XX  
 PD 04-DEC-1997.  
 XX  
 PF 30-MAY-1997; 97WO-FI00339.  
 XX  
 PR 30-MAY-1996; 96FI-0002269.  
 XX  
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.  
 XX  
 PI Heiskanen T, Lankinen H, Lundkvist A, Vahteri A;  
 XX  
 DR WPI; 1998-032807/03.  
 XX  
 PT Selection of target pathogen inhibiting substances - useful for  
 PT comparative drug design to provide therapeutically active,  
 PT protective and prophylactic substances  
 XX  
 PS Claim 16; Page 69; 81pp; English.  
 XX  
 CC This is a peptide which is useful as a leader molecule. It is used  
 CC for the development of substances active against enveloped pathogens  
 CC such as viruses from the family Bunyaviridae. Its selection  
 CC involved the reaction of a ligand library with the bound target  
 CC pathogen and the pathogen bound ligands are subjected to competitive  
 CC affinity-elution with at least one neutralising substance. The method  
 CC is useful for comparative drug design to provide therapeutically  
 CC protective and/or prophylactic substances and developing combinational  
 CC therapies as well as for pathogen diagnostics. The method also  
 CC identifies "leader molecules", which not only have great affinity,  
 CC but also mimic the functions of known neutralising substances.  
 XX  
 XX Sequence 7 AA;

Query Match 55.8%; Score 24; DB 19; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WYF 7  
   |||  
Db 2 wyf 4

Search completed: January 17, 2001, 13:47:27  
Job time: 266 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:44:46 ; Search time 13.66 Seconds  
(without alignments)  
34.795 Million cell updates/sec

Title: US-08-765-837-7  
Perfect score: 43  
Sequence: 1 ETKSWYF 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 455

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66: \*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	41.9	7	2 S21230	dermorphin (Trp-4,
2	18	41.9	7	2 PN0649	alpha-dextrin endo
3	15	34.9	4	2 PT0661	T-cell receptor be
4	15	34.9	5	2 PT0580	T-cell receptor be
5	15	34.9	6	2 A31263	dihydrofolate redu
6	15	34.9	6	2 A61068	locustakinin - mig
7	15	34.9	7	2 A61081	tryptophyllin, bas
8	15	34.9	7	2 PT0087	ribulose-bisphosph
9	14	32.6	7	2 S33567	tubulin beta-3 cha
10	13	30.2	5	2 PT0278	Ig heavy chain CRD
11	13	30.2	5	2 PT0281	Ig heavy chain CRD
12	13	30.2	7	2 A58512	venom heptapeptide
13	13	30.2	7	2 S17976	glucose isomerase
14	13	30.2	7	2 PC2132	FMRFamide-related
15	13	30.2	7	2 S33244	neuromodulatory pe
16	13	30.2	7	2 S33245	neuromodulatory pe
17	12	27.9	3	2 F37196	bradykinin-potenti
18	12	27.9	5	2 G37196	bradykinin-potenti
19	12	27.9	5	2 PT0729	T-cell receptor be
20	12	27.9	6	2 S66195	alcohol dehydrogen
21	12	27.9	6	2 B31263	dihydrofolate redu
22	12	27.9	6	2 PT0532	T-cell receptor be
23	12	27.9	6	2 PT0637	T-cell receptor be
24	12	27.9	6	2 PT0641	T-cell receptor be
25	12	27.9	6	2 PT0726	T-cell receptor be
26	12	27.9	6	2 PD0028	pev-kinin 2 - pena
27	12	27.9	7	2 PT0526	T-cell receptor be
28	12	27.9	7	2 PT0642	T-cell receptor be
29	12	27.9	7	4 A58725	virotaxin - destro

30 11 25.6 4 2 A34626 RPH-related neuro  
31 11 25.6 4 2 B53284 T-cell receptor be  
32 11 25.6 5 2 A32516 cholecystokinin-5  
33 11 25.6 5 2 B31836 20K protein - Rick  
34 11 25.6 5 2 A60803 neuropeptide - sea  
35 11 25.6 5 2 JH0253 gut pentapeptide -  
36 11 25.6 5 2 PT0308 Ig heavy chain CRD  
37 11 25.6 6 2 B34835 dnaA protein - pse  
38 11 25.6 6 2 B35640 cerebellar degener  
39 11 25.6 6 2 PT0629 T-cell receptor be  
40 11 25.6 6 2 PT0519 T-cell receptor b  
41 11 25.6 6 2 F41946 hypothetical TCU3  
42 11 25.6 6 4 I79564 triacylglycerol li  
43 11 25.6 7 2 S57274 hypothetical prote  
44 11 25.6 7 2 S09652 major fat-globule  
45 11 25.6 7 2 B48394

## ALIGNMENTS

## RESULT 1

S21230  
dermorphin (Trp-4, Asn-7) [validated] - two-colored leaf frog (fragment)  
C:Species: Phyllomedusa bicolor (two-colored leaf frog)  
C:Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 18-Aug-2000  
C:Accession: S21230

R:Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Krell, FEBS Lett. 302, 151-154, 1992  
A:Title: Identification and characterization of two dermorphins from skin extracts of  
A:Reference number: S21152; MUID:92339502

A:Accession: S21230  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MIG>

C:Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

Query Match 41.9%; Score 18; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WY 6  
Db 4 WY 5

## RESULT 2

PN0649  
alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) - Bacillus sp. (strain S-1) (f  
C:Species: Bacillus sp.  
C:Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 07-May-1999  
C:Accession: PN0649

R:Kim, C.H.; Choi, H.I.; Lee, D.S.  
Biosci. Biotechnol. Biochem. 57, 1632-1637, 1993  
A:Title: Purification and biochemical properties of an alkaline pullulanase from alka  
A:Reference number: PN0649; MUID:94080025  
A:Accession: PN0649

A:Molecule type: protein  
A:Residues: 1-7 <KIM>

C:Comment: This enzyme is used together with glucoamylase to improve the efficiency of  
ntent in high maltose syrups

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 41.9%; Score 18; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SWY 6  
Db 5 SWF 7

## RESULT 3

PT0661

T-cell receptor beta chain V-D-J region (121-1BV) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PT0661

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0661

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-4 &lt;FE&gt;

A:Experimental source: day 4 postnatal thymus, strain BALB/c

C:Keywords: T-cell receptor

## Query Match

34.9%; Score 15; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SW 5

||

Db 2 SW 3

## RESULT 4

PT0580

T-cell receptor beta chain V-D-J region (159-2B) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PT0580

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0580

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 &lt;FE&gt;

A:Experimental source: day 19 fetal thymus, strain BALB/c

C:Keywords: T-cell receptor

## Query Match

34.9%; Score 15; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SW 5

||

Db 3 SW 4

## RESULT 5

A31263

dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - Plasmodium f

C:Species: Plasmodium falciparum

C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996

C:Accession: A31263

R:Peterson, D.S.; Walliker, D.; Wellens, T.E.

Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988

A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase

A:Reference number: A94217; MUID:89057886

A:Accession: A31263

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-6 &lt;PET&gt;

C:Keywords: methyltransferase; NADP; oxidoreductase

## Query Match

34.9%; Score 15; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SW 5

||

Db 2 SW 3

## RESULT 6

A61068

locustakinin - migratory locust

C:Species: Locusta migratoria (migratory locust)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 08-Dec-1995

C:Accession: A61068

R:Schoofs, L.; Holman, G.M.; Proost, P.; Van Damme, J.; Hayes, T.K.; De Loof, A.

Regul. Pept. 37, 49-57, 1992

A:Title: Locustakinin, a novel myotropic peptide from Locusta migratoria, isolation,

A:Reference number: A61068; MUID:92262851

A:Accession: A61068

A:Molecule type: protein

A:Residues: 1-6 &lt;SCH&gt;

C:Keywords: amidated carboxyl end; cephalomyotropic peptide; neuropeptide

F:6/Modified site: amidated carboxyl end (Gly) #status experimental

## Query Match

34.9%; Score 15; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SW 5

||

Db 4 SW 5

## RESULT 7

A61081

tryptophyllin, basic - Rohde's leaf frog

C:Species: Phyllomedusa rohdei (Rohde's leaf frog)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Aug-2000

C:Accession: A61081

R:Montecucchi, P.C.; Vincenti, M.; Lazzarini, A.M.; Rusconi, L.; Erspamer, V.

Int. J. Pept. Protein Res. 33, 391-395, 1989

A:Title: Isolation, structure determination and synthesis of a novel tryptophan-conta

A:Reference number: A61081

A:Accession: A61081

A:Molecule type: protein

A:Residues: 1-7 &lt;MON&gt;

C:Comment: The biological activity of this peptide was not determined.

C:Superfamily: unassigned animal peptides

C:Keywords: amidated carboxyl end; hydroxyproline; skin

F:3/Modified site: 4-hydroxyproline (Pro) #status experimental

F:7/Modified site: amidated carboxyl end (Pro) #status experimental

## Query Match

34.9%; Score 15; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SW 5

||

Db 4 SW 5

## RESULT 8

PT0087

ribulose-bisphosphate carboxylase (EC 4.1.1.39) large chain - Arabidopsis thaliana (f

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 10-Apr-1996 #sequence\_revision 24-May-1996 #text\_change 24-Nov-1999

C:Accession: PT0087

R:Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.

submitted to JIPID, December 1995

A:Description: Two dimensional electrophoresis of plant proteins and standardization

A:Reference number: PN0173

A:Accession: PT0087  
A:Molecule type: protein  
A:Residues: 1-7 <TSU>

A:Experimental source: leaf  
C:Keywords: acetylated amino end; carbon-carbon lyase; carboxy-lyase  
F:1/Modified site: acetylated amino end (Ser) #status experimental

Query Match 34.9%; Score 15; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETK 3  
|||  
DB 5 ETK 7

RESULT 9  
S33567  
tubulin beta-3 chain - fruit fly (Drosophila melanogaster) (fragment)  
C:Species: Drosophila melanogaster  
C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 21-Jul-2000  
C:Accession: S33567  
R:Hintz, U.; Wolk, A.; Renkawitz-Pohl, R.  
Development 116, 543-554, 1992  
A:Title: Ultrathorax is a regulator of beta-3 tubulin expression in the Drosophila vis  
A:Reference number: S33567; MUID:93170162  
A:Accession: S33567  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-7 <HIN>  
A:Cross-references: EMBL:X68393; NID:g7669; PIDN:CAA48459.1; PID:g4377448  
C:Genetics:  
A:Gene: FlyBase:fbgr  
A:Cross-references: FlyBase:FBgn0003888  
A:Introns: 5/3

Query Match 32.6%; Score 14; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KSW 5  
|||  
DB 5 KFW 7

RESULT 10  
PT0278  
Ig heavy chain CRD3 region (clone 4-88) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0278  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
A:Reference number: PT0222; MUID:91108337  
A:Accession: PT0278  
A:Molecule type: DNA  
A:Residues: 1-5 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 30.2%; Score 13; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7  
|||  
DB 1 YF 2

RESULT 11  
PT0281

Ig heavy chain CRD3 region (clone 4-91C) - human (fragment)  
C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0281

R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an  
A:Reference number: PT0222; MUID:91108337

A:Accession: PT0281

A:Molecule type: DNA

A:Residues: 1-5 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 30.2%; Score 13; DB 2; Length 5;  
Best Local Similarity 33.3%; Pred. No. 2e+05;  
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KSW 5  
::|  
DB 2 ENW 4

RESULT 12  
A58512  
venom heptapeptide - cone shell (Conus imperialis)  
C:Species: Conus imperialis (imperial cone)  
C:Date: 19-Mar-1997 #sequence\_revision 11-Apr-1997 #text\_change 07-May-1999  
C:Accession: A58512  
R:Craig, A.G.; Jimenez, E.C.; Dykert, J.; Nielsen, D.B.; Gulyas, J.; Abogadie, F.C.;  
J. Biol. Chem. 272, 4689-4698, 1997  
A:Title: A novel post-translational modification involving bromination of tryptophan.  
A:Reference number: A58512; MUID:97184108  
A:Accession: A58512  
A:Molecule type: protein  
A:Residues: 1-7 <CRA>  
C:Superfamily: unassigned conotoxins  
C:Keywords: amidated carboxyl end; bromine; pyroglutamic acid; venom  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:6/Modified site: 6-bromotryptophan (Trp) #status experimental  
F:7/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 30.2%; Score 13; DB 2; Length 7;  
Best Local Similarity 33.3%; Pred. No. 2e+05;  
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KSW 5  
::|  
DB 4 QAW 6

RESULT 13  
S17976  
glucose isomerase - Thermoanaerobacterium saccharolyticum (fragment)  
C:Species: Thermoanaerobacterium saccharolyticum  
C:Date: 12-Feb-1998 #sequence\_revision 12-Feb-1998 #text\_change 17-Apr-1998  
C:Accession: S17976  
R:Lee, C.; Zeikus, J.G.  
Biochem. J. 273, 565-571, 1991  
A:Title: Purification and characterization of thermostable glucose isomerase from Clo  
A:Reference number: S15119; MUID:91144536  
A:Accession: S17976  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <LEE>

A:Note: the source is designated as Thermoanaerobacter strain B6A

Query Match 30.2%; Score 13; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7  
| |  
Db 4 YF 5

## RESULT 14

PC2132  
FMRamide-related heptapeptide - Panagrellus redivivus  
C:Species: Panagrellus redivivus  
C:Date: 03-May-1994 #sequence\_revision 15-Oct-1994 #text\_change 11-Jul-1997  
C:Accession: PC2132  
R:Maule, A.G.; Shaw, C.; Bowman, J.W.; Halton, D.W.; Thompson, D.P.; Geary, T.G.; Thim,  
Biochem. Biophys. Res. Commun. 200, 973-980, 1994  
A:Title: KSAVMRFamide: a novel FMRamide-related heptapeptide from the free-living nematode  
A:Reference number: PC2132; MUID:94235053  
A:Accession: PC2132  
A:Molecule type: protein  
A:Residues: 1-7 <MAV>  
C:Keywords: amidated carboxyl end  
F:7/Modified site: amidated carboxyl end (Phe) #status predicted

Query Match 30.2%; Score 13; DB 2; Length 7;  
Best Local Similarity 75.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KSWY 6  
| | |  
Db 1 KSAY 4

## RESULT 15

S33244  
neuromodulatory peptide Wamide-1 - giant African snail  
C:Species: Achatina fulica (giant African snail)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
C:Accession: S33244  
R:Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.  
FEBS Lett. 323, 104-108, 1993  
A:Title: Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of the  
A:Reference number: S33244; MUID:93265912  
A:Accession: S33244  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MIN>

Query Match 30.2%; Score 13; DB 2; Length 7;  
Best Local Similarity 40.0%; Pred. No. 2e+05;  
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETKSW 5  
| |  
Db 3 EMSVW 7

Search completed: January 17, 2001, 13:48:00  
Job time: 194 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 17, 2001, 13:47:31 ; Search time 8.41 Seconds  
(without alignments)  
26.880 Million cell updates/sec

Title: US-08-765-837-7

Perfect score: 43

Sequence: 1 ETKSWVF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 72

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	34.9	6	LOK1_LOCFMI	P41491 locusta mig
2	13	30.2	7	FAR3_HAECO	P81298 haemochus
3	13	30.2	7	FAR3_PANRE	P41874 panagrellus
4	13	30.2	7	WMA1_ACHFU	P35919 achatina fu
5	13	30.2	7	WMA3_ACHFU	P35921 achatina fu
6	12	27.9	5	BP77_BOTIN	P30425 bothrops in
7	11	25.6	5	UF01_MOUSE	P38639 mus musculus
8	11	25.6	7	WMA2_ACHFU	P35920 achatina fu
9	10	23.3	4	TUFT_HUMAN	P01858 homo sapien
10	10	23.3	5	PAP2_PARMA	P81864 pardachirus
11	10	23.3	7	IGAO_DACDE	P06294 dactylum d
12	9	20.9	5	UC22_MAIZE	P80628 zea mays (m
13	9	20.9	6	FARP_MONEX	P41966 monlexia ex
14	9	20.9	7	ALL7_CYDPO	P82158 cydia pomon
15	9	20.9	7	UC24_MAIZE	P80630 zea mays (m
16	8	18.6	7	ALL2_CARMA	P81805 carcinus ma
17	8	18.6	7	FAR2_ASCSU	P31890 ascaris suu
18	7	16.3	4	FAR3_HIRME	P42562 hirudo medi
19	7	16.3	4	FAR4_HIRME	P42563 hirudo medi
20	7	16.3	5	ALL4_CARMA	P81817 carcinus ma
21	7	16.3	5	FARP_ARTTR	P41853 artiopesthi
22	7	16.3	5	PACT_PERAM	P01373 periplaneta
23	7	16.3	6	OMV_LEPDE	P42985 leptonotars
24	7	16.3	7	ALL3_CARMA	P81806 carcinus ma
25	7	16.3	7	ALL4_CARMA	P81807 carcinus ma
26	7	16.3	7	ALL5_CARMA	P81808 carcinus ma
27	7	16.3	7	FAR5_HIRME	P42564 hirudo medi
28	7	16.3	7	GFRP_MOUSE	P99025 mus musculus
29	7	16.3	7	H77_PIG	P01153 sus scrofa
30	7	16.3	7	MNP1_LEPDE	P42984 leptonotars
31	7	16.3	7	UN06_PINPS	P81675 pinus pinas
32	6	14.0	4	ACH1_ACHFU	P35904 achatina fu
33	6	14.0	4	DCNS_PSECH	P19918 pseudomonas

34 6 14.0 4 1 FLRF\_HIRME P42561 hirudo medi  
35 6 14.0 4 1 FMRE\_MACNI P01162 macrocallis  
36 6 14.0 5 1 SUGA\_ACHDO P19991 acheta dome  
37 6 14.0 5 1 TPIS\_CANFA P54714 canis famli  
38 6 14.0 5 1 TRM3\_ECOLI P19973 escherichia  
39 6 14.0 6 1 CIP1\_MVTED P13736 mytilus edu  
40 6 14.0 6 1 CIP2\_MVTED P13737 mytilus edu  
41 6 14.0 6 1 VP19\_HSVIK P23210 herpes simp  
42 6 14.0 7 1 CCF1\_ENTFA P20104 enterococcus  
43 6 14.0 7 1 CIA\_ENTFA P19332 enterococcus  
44 6 14.0 7 1 FAR1\_ASCSU P31889 ascaris suu  
45 6 14.0 7 1 FAR1\_HELTI P41871 helisma tr

#### ALIGNMENTS

RESULT 1  
LOK1\_LOCFMI  
ID LOK1\_LOCFMI STANDARD; PRT; 6 AA.  
AC P41491;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE LOCUSTAKININ I.  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
OC Acridomorpha; Acridoidea; Acrididae; Locusta.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=CORPORA CARDIACA;  
RX MEDLINE=92262851; PubMed=1585017;  
RA Schoofs L., Holman G.W., Proost P., van Damme J., Hayes T.K.,  
RA de Loof A.;  
RT "Locustakinin, a novel myotropic peptide from Locusta migratoria,  
RT isolation, primary structure and synthesis.";  
RL Regul. Pept. 37:49-57(1992).  
CC -!- FUNCTION: MYOTROPIC PEPTIDE. MAY BE IMPORTANT IN THE STIMULATION  
CC OF ION TRANSPORT AND INHIBITION OF DIURETIC ACTIVITY IN MALPIGHIAN  
CC TUBULES.  
CC PIR; A61068; A61068.  
DR Neuropeptide; Amidation.  
KW MOD\_RES 6  
FT SEQUENCE 6 AA; 654 MW; 586365A5B9CDB000 CRC64;  
SQ

Query Match 34.9%; Score 15; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SW 5  
Db 4 SW 5

RESULT 2  
FAR3\_HAECO  
ID FAR3\_HAECO STANDARD; PRT; 7 AA.  
AC P81298;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE FMRFAMIDE-LIKE NEUROPEPTIDE PF3 (KSAYMRF-AMIDE).  
OS Haemochus contortus.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Trichostrongyloidea; Haemonchidae; Haemonchus.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=NEURON;  
RA Marks N.J., Sangster N.C., Maule A.G., Halton D.W., Geary T.G.,  
RA Thompson D.P., Shaw C.;  
RL Submitted (MAY-1998) to the SWISS-PROT data bank.

CC -1- FUNCTION: ACTIVE ON NEUROMUSCULATURE.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;

Query Match 30.2%; Score 13; DB 1; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KSWY 6  
 | | |  
 Db 1 KSAI 4

RESULT 3  
 FAR3\_PANRE STANDARD; PRT; 7 AA.  
 ID FAR3\_PANRE  
 AC P41874;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE PF3 (KSAYMRF-AMIDE).  
 OS Panagrellus redivivus.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RX MEDLINE=94235053; PubMed=8179635;  
 RA Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P.,  
 RA Geary T.G., Thim L.;  
 RT "KSAYMRFamide: a novel FMRFamide-related heptapeptide from the free-living nematode, Panagrellus redivivus, which is myoactive in the parasitic nematode, Ascaris suum.";  
 RL Biochem. Biophys. Res. Commun. 200:973-980(1994).  
 CC -1- FUNCTION: MYOACTIVE; INDUCES A RAPID CONCENTRATION-DEPENDENT MUSCLE TENSION INCREASE.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR PIR; PC2132.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 902 MW; 69D4068B5DC5B350 CRC64;

Query Match 30.2%; Score 13; DB 1; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KSWY 6  
 | | |  
 Db 1 KSAI 4

RESULT 4  
 WWAL\_ACHFU STANDARD; PRT; 7 AA.  
 ID WWAL\_ACHFU  
 AC P35919;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE WWAMIDE-1.  
 OS Achatina fulica (Giant African snail).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 OC Achatinacea; Achatinidae; Achatina.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=GANGLION;  
 RX MEDLINE=93265912; PubMed=8495720;  
 RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;  
 RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from

RT ganglia of the African giant snail, Achatina fulica.";  
 RL FEBS Lett. 323:104-108(1993).  
 CC -1- FUNCTION: EXHIBITS MODULATORY EFFECTS ON THE PERIPHERAL NERVOUS  
 CC SYSTEM. INHIBITS ACTIVITY ON A CENTRAL NEURON.  
 DR PIR; S33245; S33245.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 993 MW; 7362D5B69B041310 CRC64;

Query Match 30.2%; Score 13; DB 1; Length 7;  
 Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ETKSW 5  
 | | |  
 Db 3 EMSVW 7

RESULT 5  
 WWA3\_ACHFU STANDARD; PRT; 7 AA.  
 ID WWA3\_ACHFU  
 AC P35921;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE WWAMIDE-3.  
 OS Achatina fulica (Giant African snail).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 OC Achatinacea; Achatinidae; Achatina.  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=GANGLION;  
 RX MEDLINE=93265912; PubMed=8495720;  
 RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;  
 RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of the African giant snail, Achatina fulica.";  
 RL FEBS Lett. 323:104-108(1993).  
 DR PIR; S33244; S33244.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 965 MW; 7362D5B69B132310 CRC64;

Query Match 30.2%; Score 13; DB 1; Length 7;  
 Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ETKSW 5  
 | | |  
 Db 3 EMSVW 7

RESULT 6  
 BPP7\_BOTIN STANDARD; PRT; 5 AA.  
 ID BPP7\_BOTIN  
 AC P30425;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE BRADYKININ-POTENTIATING PEPTIDE S5,2 (5A) (ANGIOTENSIN-CONVERTING ENZYME INHIBITOR).  
 OS Bothrops insularis (Island jararaca) (Queimada jararaca).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 OC Viperidae; Crotalinae; Bothrops.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=VENOM;  
 RX MEDLINE=90351557; PubMed=2386615;  
 RA Cintra A.C.O., Vieira C.A., Giglio J.R.;

RT "Primary structure and biological activity of bradykinin potentiating peptides from Bothrops insularis snake venom.";



RL J. Protein Chem. 9:221-227(1990).  
 CC -!- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE  
 CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF THE  
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.  
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.

DR PIR: G37196; G37196.

KW Hypotensive agent; Venoim.

FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

SQ SEQUENCE 5 AA; 629 MW; 776DC37326B00000 CRC64;

Query Match 27.9%; Score 12; DB 1; Length 5;  
 Best Local Similarity 33.3%; Pred. No. 8.9e+04;  
 Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 KSW 5  
 DB 1 OKW 3

RESULT 7  
 UF01\_MOUSE STANDARD; PRT; 5 AA.  
 ID UF01\_MOUSE  
 AC P38639;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF FIBROBLASTS (P19) (FRAGMENT).

OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]

RP SEQUENCE.

RC TISSUE=FIBROBLAST;

RX MEDLINE=95009907; PubMed=7523108;

RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;

RT "Separation and sequencing of familial and novel murine proteins

using preparative two-dimensional gel electrophoresis.";

RL Electrophoresis 15:735-745(1994).

CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN

NON\_TER 5 5

FT PROTEIN IS: 6.6, ITS MW IS: 19 KDA.

SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;  
 Query Match 25.6%; Score 11; DB 1; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 W 5  
 DB 1 W 1

RESULT 8  
 WWA2\_ACHFU STANDARD; PRT; 7 AA.  
 ID WWA2\_ACHFU  
 AC P35920;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE WWA2\_ACHFU-2.  
 OS Achatina fulica (Giant African snail).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 CC Achatinacea; Achatinidae; Achatina.  
 RN [1]

RP SEQUENCE.

RC TISSUE=GANGLION;

RX MEDLINE=93265912; PubMed=8495720;

RA Minakata H., Ikeda T., Munaka Y., Kobayashi M., Nomoto K.;

RT "Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from

ganglia of the African giant snail, Achatina fulica.";

RL FEBS Lett. 323:104-108(1993).

KW Neuropeptide; Amidation.  
 FT MOD\_RES 7 7 AMIDATION.  
 SQ SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 25.6%; Score 11; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 W 5  
 DB 1 W 1

RESULT 9  
 TUFT\_HUMAN STANDARD; PRT; 4 AA.  
 ID TUFT\_HUMAN  
 AC P01858;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE PHAGOCYTOSIS-STIMULATING PEPTIDE (TUFTSIN).  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]

RP SEQUENCE.

RX MEDLINE=72187087; PubMed=4112769;

RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;

RT "The characteristics, isolation and synthesis of the phagocytosis

stimulating peptide tuftsin.";

RL Biochem. Biophys. Res. Commun. 47:172-179(1972).

RN [2]

RP IMMUNOGLOBULIN CLASS.

RX MEDLINE=68091045; PubMed=4169272;

RA Fidalgo B.V., Najjar V.A.;

RT "The physiological role of the lymphoid system. VI. The stimulatory

effect of leucophilic gamma globulin (leucokinin) on the phagocytic

activity of human polymorphonuclear leucocyte.";

RL Biochemistry 6:3386-3392(1967).

CC -!- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSIBLY TO THE

CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE

MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN.

CC TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC

ACTIVITY OF NEUTROPHILS.

DR PIR: A02147; A02147.

DR MIN: 191150; -.

SQ SEQUENCE 4 AA; 501 MW; 74176321C00000000 CRC64;

Query Match 23.3%; Score 10; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TK 3  
 DB 1 TK 2

RESULT 10  
 PAP2\_PARMA STANDARD; PRT; 5 AA.  
 ID PAP2\_PARMA  
 AC P81864;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PARDAXIN II (PXII) (FRAGMENT).  
 OS Pardachirus marmoratus (Red sea mores sole).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
 CC Soleiidei; Soleidae; Pardachirus.  
 RN [1]

RP SEQUENCE.  
 RC TISSUE=SKIN SECRETION;  
 RX MEDLINE=87057369; PubMed=3782138;  
 RA Lazarovici P., Primor N., Loew L.M.;  
 RT "Purification and pore-forming activity of two hydrophobic  
 RT polypeptides from the secretion of the Red sea moses sole (Pardachirus  
 RT marmoratus).";  
 RL J. Biol. Chem. 261:16704-16713(1986).  
 CC -1- FUNCTION: EXHIBITS UNUSUAL SHARK REPELLENT AND SURFACTANT  
 CC PROPERTIES. FORMS VOLTAGE-DEPENDENT, ION-PERMEABLE CHANNELS  
 CC IN MEMBRANES. AT HIGH CONCENTRATION CAUSES CELL MEMBRANE LYSIS.  
 CC -1- SUBUNIT: MONOMER. IN AQUEOUS SOLUTION EXISTS AS A TETRAMER.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.  
 KW Toxin.  
 FT NON\_TER 5 5  
 SQ SEQUENCE 5 AA; 614 MW; 7769C9C8100000 CRC64;

Query Match 23.3%; Score 10; DB 1; Length 5;  
 Best Local Similarity 33.3%; Pred. No. 8.9e+04;  
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7  
 : :  
 DB 2 FFF 4

RESULT 11  
 IGAO\_DACDE STANDARD; PRT; 7 AA.  
 ID IGAO\_DACDE  
 AC P06294;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE GALACTOSE OXIDASE INHIBITOR.  
 OS Dactylium dendroides (Cladobotryum dendroides).  
 OC Eukaryota; Fungi; Ascomycota; Hypocreales; Hypocreaceae; Hypomyces.  
 RN [1]  
 RP SEQUENCE.  
 RA Avigad G., Markus Z.;  
 RT "Identification of a peptide inhibitor of galactose oxidase from  
 RT Dactylium dendroides.";  
 RL Fed. Proc. 31:447-447(1972).  
 CC -1- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE  
 CC GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY  
 CC BINDING TO ITS PROSTHETIC COPPER GROUP.  
 DR PIR; A01341; XEYDGD.  
 KW Copper; Metalloenzyme inhibitor.  
 SQ SEQUENCE 7 AA; 706 MW; 75BB01A456D87DB0 CRC64;

Query Match 23.3%; Score 10; DB 1; Length 7;  
 Best Local Similarity 56.7%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TKS 4  
 : :  
 DB 5 TES 7

RESULT 12  
 UC22\_MAIZE STANDARD; PRT; 5 AA.  
 ID UC22\_MAIZE  
 AC P80628;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 474)  
 DE (FRAGMENT).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.

RN SEQUENCE.  
 RP TISSUE=COLEOPTILE;  
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RT Pernollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize two dimensional gel protein database: towards an integrated  
 RT genome analysis program.";  
 RL Theor. Appl. Genet. 93:997-1005(1996).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.1, ITS MW IS: 30.4 KDA.  
 DR MAIZE-2DPAGE; P80628; COLEOPTILE.  
 FT MAIZEDB; 123954; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 5 5  
 SQ SEQUENCE 5 AA; 654 MW; 72CB19C9C0300000 CRC64;

Query Match 20.9%; Score 9; DB 1; Length 5;  
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7  
 : :  
 DB 2 FF 3

RESULT 13  
 FARP\_MONEX STANDARD; PRT; 6 AA.  
 ID FARP\_MONEX  
 AC P41966;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE CNFRF-AMIDE.  
 OS Moniezia expansa (Sheep tapeworm).  
 OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Cestoda; Eucestoda;  
 OC Cyclophyllidae; Anoplocephalidae; Moniezia.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93312289; PubMed=8323531;  
 RA Maule A.G., Shaw C., Hulton D.W., Thim L.;  
 RT "GNFRFamide: a novel FMRFamide-immunoreactive peptide isolated from  
 RT the sheep tapeworm, Moniezia expansa.";  
 RL Biochem. Biophys. Res. Commun. 193:1054-1060(1993).  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 6 6  
 SQ SEQUENCE 6 AA; 787 MW; 69D409C9C4481000 CRC64;

Query Match 20.9%; Score 9; DB 1; Length 6;  
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7  
 : :  
 DB 3 FF 4

RESULT 14  
 ALL7\_CYPDO STANDARD; PRT; 7 AA.  
 ID ALL7\_CYPDO  
 AC P82158;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYDIASTATIN 7.  
 OS Cydia pomonella (Codling moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Tortricoidea; Tortricidae; Olethreutinae; Cydia.  
 RN [1]

```
RP SEQUENCE..
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
FT SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;

Query Match 20.9%; Score 9; DB 1; Length 7;
Best Local Similarity 40.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KSWYF 7
Db 1 KMYDF 5

RESULT 15
UC24_MAIZE STANDARD; PRT; 7 AA.
ID UC24_MAIZE
AC P80630;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 447)
DE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RP SEQUENCE.
RC TISSUE=COLEOPTILE;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -|- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.0, ITS MW IS: 30.0 KDA.
DR MAIZE-2DPAGE: P80630; COLEOPTILE.
DR MAIZEDB; 123956; -.
FT NON_TER 1
FT NON_TER 7
FT SEQUENCE 7 AA; 665 MW; 6DC1B5B33DC1B5D0 CRC64;
```

```
Query Match 20.9%; Score 9; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KS 4
Db 4 KS 5
```

Search completed: January 17, 2001, 13:49:14  
Job time: 103 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 17, 2001, 13:47:16 ; Search time 19.2 Seconds  
(without alignments)  
42.732 Million cell updates/sec

Title: US-08-765-837-7  
Perfect score: 43  
Sequence: 1 ETKSWYF 7

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 66

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	32.6	7	P82065	P82065 litoria rub
2	11	25.6	6	P82541	P82541 spinacia ol
3	11	25.6	6	P82096	P82096 litoria rub
4	11	25.6	7	Q95945	Q95945 saccharomyc
5	11	25.6	7	O49223	O49223 glycine max
6	10	23.3	6	P82181	P82181 spinacia ol
7	10	23.3	6	P82182	P82182 spinacia ol
8	10	23.3	7	Q47029	Q47029 enterobacte
9	9	20.9	5	P82070	P82070 litoria rub
10	9	20.9	5	P82071	P82071 litoria rub
11	9	20.9	5	P82072	P82072 litoria rub
12	9	20.9	5	P82073	P82073 litoria rub
13	8	18.6	7	Q65578	Q65578 bovine herp
14	7	16.3	7	P70804	P70804 azotobacter
15	7	16.3	7	O50556	O50556 actinobacil
16	7	16.3	7	Q15903	Q15903 homo sapien
17	7	16.3	7	O99182	O99182 pterolebias
18	7	16.3	7	O9YVE3	O9YVE3 human adeno
19	7	16.3	7	Q9YQ10	Q9YQ10 porcine tra

20	7	16.3	7	12	Q9YIR0	Q9YIR0 human adeno
21	7	16.3	7	12	Q9YIQ9	Q9YIQ9 human adeno
22	7	16.3	7	13	P82101	P82101 litoria rub
23	6	14.0	5	13	P82099	P82099 litoria rub
24	6	14.0	5	13	P82100	P82100 litoria rub
25	6	14.0	6	4	Q08720	Q08720 homo sapien
26	6	14.0	7	2	Q54248	Q54248 streptomyce
27	6	14.0	7	2	O07354	O07354 synechococc
28	6	14.0	7	2	O34028	O34028 pseudomonas
29	6	14.0	7	4	Q15897	Q15897 homo sapien
30	6	14.0	7	5	Q9YVN9	Q9YVN9 drosophila
31	6	14.0	7	8	O98866	O98866 spinacia ol
32	6	14.0	7	10	P82445	P82445 nicotiana t
33	5	11.6	4	11	Q08433	Q08433 rattus norv
34	5	11.6	5	10	Q99007	Q99007 hordeum vul
35	5	11.6	7	2	P72081	P72081 nocardia la
36	5	11.6	7	2	Q47505	Q47505 escherichia
37	5	11.6	7	2	Q47477	Q47477 escherichia
38	5	11.6	7	6	Q28742	Q28742 oryctolagus
39	5	11.6	7	8	P92421	P92421 psathyrosta
40	5	11.6	7	8	P92385	P92385 hordeum mar
41	5	11.6	7	8	P92210	P92210 agropyron c
42	5	11.6	7	8	P92214	P92214 amblyopyrum
43	5	11.6	7	8	P92218	P92218 australopyr
44	5	11.6	7	8	P92221	P92221 bromus iner
45	5	11.6	7	12	Q66205	Q66205 porcine tra

#### ALIGNMENTS

RESULT 1

P82065	PRELIMINARY:	PRT:	7 AA.
ID	P82065		
AC	P82065;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)		
DE	TRPTOPHYLLIN 5.1.		
OS	Litoria rubella (Desert tree frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;		
OC	Litoria.		
OX	NCBI_TaxID=104895;		
RN	[1]		
RP	SEQUENCE, AND MASS SPECTROMETRY.		
RC	TISSUE=SKIN SECRETION;		
RA	Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Cao C.,		
RA	Tyler M.J., Wallace J.C.;		
RT	"The structure of new peptides from the Australian red tree frog		
RT	'Litoria rubella'. The skin peptide profile as a probe for the study		
RT	of evolutionary trends of amphibians.";		
RL	Aust. J. Chem. 49:955-963(1996).		
CC	!- FUNCTION: MAY ACT AS A NEUROMODULATOR OR NEUROTRANSMITTER.		
CC	!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.		
CC	!- MASS SPECTROMETRY: MW=965; METHOD=FAB.		
KW	Amphibian skin; Amidation; Neuropeptide.		
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.		
FT	MOD_RES 7 7 AMIDATION.		
SQ	SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;		

Query Match 32.6%; Score 14; DB 13; Length 7;  
Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 WY 6

Db 4 WF 5

RESULT 2

P82541

ID P82541 PRELIMINARY; PRT; 6 AA.  
AC P82541;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S19 BETA (FRAGMENT).  
OS Spinacia oleracea (Spinach).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;  
OC Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV. ALVARO;  
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;  
RT Identification of all the proteins in the small subunit of an  
RT organelle (chloroplast) ribosome.";  
RL J. Biol. Chem. 0:0-0(2000).  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -1- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PL. S19 BETA  
CC FORM IS THE MINOR BASIC FORM.  
CC -1- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.  
DR INTERPRO: IPR002222; -  
DR PROSITE: PS00323; RIBOSOMAL\_S19; PARTIAL.  
KW Ribosomal protein; Chloroplast.  
FT NON\_TER 6  
SQ SEQUENCE 6 AA; 732 MW; 63333735A11C000 CRC64;

Query Match 25.6%; Score 11; DB 10; Length 6;  
Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 TKS 4  
I: I  
D: I  
DB 1 TRS 3

RESULT 3  
P82096 PRELIMINARY; PRT; 6 AA.  
ID P82096  
AC P82096;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE ELECTRIN 1.  
OS Litoria rubella (Desert tree frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
OC Litoria.  
OX NCBI\_TaxID=104895;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=SKIN SECRETION;  
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;  
RA "Peptides from the skin glands of the Australian buzzing tree frog  
RT Litoria electrica. Comparison with the skin peptides from Litoria  
RT rubella.";  
RL Aust. J. Chem. 52:0-0(1999).  
KW Amphibian skin; Amidation.  
FT MOD\_RES 6  
SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A000 CRC64;

Query Match 25.6%; Score 11; DB 13; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 W 5  
I: I  
D: I  
DB 5 W 5

RESULT 4  
Q95945 PRELIMINARY; PRT; 7 AA.  
ID Q95945  
AC Q95945;  
DT 01-FEB-1997 (TReMBLrel. 02, Created)  
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE INSIDE INTRON 5 (FRAGMENT).  
OS Saccharomyces cerevisiae (Baker's yeast).  
OG Mitochondrion.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8106985; PubMed=6254986;  
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;  
RT "Assembly of the mitochondrial membrane system. Structure and  
RT nucleotide sequence of the gene coding for subunit 1 of yeast  
RT cytochrome oxidase.";  
RL J. Biol. Chem. 255:11927-11941(1980).  
DR EMBL; V00694; CAA24066.1; -  
KW Mitochondrion.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match 25.6%; Score 11; DB 8; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 W 5  
I: I  
D: I  
DB 4 W 4

RESULT 5  
O49223 PRELIMINARY; PRT; 7 AA.  
ID O49223  
AC O49223;  
DT 01-JUN-1998 (TReMBLrel. 06, Created)  
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE HMG-1-LIKE PROTEIN (FRAGMENT).  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
OC Fabales; Fabaceae; Papilionoideae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. ESSEX; TISSUE=ROOTS;  
RX MEDLINE=91367679; PubMed=1891369;  
RA Laux T., Goldberg R.B.;  
RT "A plant DNA binding protein shares highly conserved sequence motifs  
RT with HMG-box proteins.";  
RL Nucleic Acids Res. 19:4769-4769(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. ESSEX; TISSUE=ROOTS;  
RA Mahalingam R., Knap H.T.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF047050; AAC03556.1; -  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;

Query Match 25.6%; Score 11; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 5 W 5
Db 2 W 2

RESULT 6
P82181 ID P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 BETA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALVARO; TISSUE=LEAF;
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT "Identification of all the proteins in the large subunit of an
organelle (chloroplast) ribosome.";
RL J. Biol. Chem. 0:0-0(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- SIMILARITY: IPR002363; -.
DR INTERPRO: IPR002363; -.
KW PROSITE; PS01109; RIBOSOMAL L10; 1.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 23.3%; Score 10; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TK 3
Db 5 TK 6

RESULT 7
P82182 ID P82182 PRELIMINARY; PRT; 6 AA.
AC P82182;
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 GAMMA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALVARO; TISSUE=LEAF;
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT "Identification of all the proteins in the large subunit of an
organelle (chloroplast) ribosome.";
RL J. Biol. Chem. 0:0-0(2000).
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -!- SIMILARITY: IPR002363; -.
DR INTERPRO: IPR002363; -.
KW PROSITE; PS01109; RIBOSOMAL L10; 1.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 23.3%; Score 10; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TK 3
Db 5 TK 6

RESULT 8
P82183 ID P82183 PRELIMINARY; PRT; 7 AA.
AC P82183;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE AAD A1 PROTEIN (FRAGMENT).
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94079349; PubMed-8257126;
RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT "Analysis of the aac(3)-Via gene encoding a novel 3'-N-
acetyltransferase.";
RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR EMBL; M88012; AAA16193.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 23.3%; Score 10; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TK 3
Db 2 TK 3

RESULT 9
P82070 ID P82070 PRELIMINARY; PRT; 5 AA.
AC P82070;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELLIDIN 1.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=SKIN SECRETION;
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowle J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
'Litoria rubella', the skin peptide profile as a probe for the study
of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
ANTIBIOTIC ACTIVITY.
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -!- MASS SPECTROMETRY: MW=598; METHOD=FAB.
KW Amphibian skin.
```

```
SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;

Query Match 20.9%; Score 9; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 3.7e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7
   :|
Db 3 FF 4

RESULT 10
P82071 ID P82071 PRELIMINARY; PRT; 5 AA.
AC P82071;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE RUBELLIDIN 2.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION;
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
   Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
   'Litoria rubella', the skin peptide profile as a probe for the study
   of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=626; METHOD=FAB.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;

Query Match 20.9%; Score 9; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 3.7e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7
   :|
Db 3 FF 4

RESULT 11
P82072 ID P82072 PRELIMINARY; PRT; 5 AA.
AC P82072;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE RUBELLIDIN 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION;
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
   Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
   'Litoria rubella', the skin peptide profile as a probe for the study
   of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
```

```
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW Amphibian skin; Amidation.
FT MOD_RES 5 5 AMIDATION.
SQ SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

Query Match 20.9%; Score 9; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 3.7e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7
   :|
Db 3 FF 4

RESULT 12
P82073 ID P82073 PRELIMINARY; PRT; 5 AA.
AC P82073;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE RUBELLIDIN 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE-SKIN SECRETION;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
   Litoria electrica. Comparison with the skin peptides from Litoria
   rubella.";
RL Aust. J. Chem. 52:0-0(1999).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;

Query Match 20.9%; Score 9; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 3.7e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YF 7
   :|
Db 3 FF 4

RESULT 13
P85578 ID P85578 PRELIMINARY; PRT; 7 AA.
AC P85578;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE HYPOTHETICAL PROTEIN (FRAGMENT).
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COOPER;
RX MEDLINE=95313343; PubMed=7793062;
RA Vicek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
   Letchworth G.J., Schwzyer M.;
```



RT "Nucleotide sequence analysis of a 30-kb region of the bovine  
 RT herpesvirus 1 genome which exhibits a colinear gene arrangement with  
 RT the UL21 to UL4 genes of herpes simplex virus.";  
 RL Virology 210:100-108(1995).  
 DR EMBL; Z48053; CAA88130.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 758 MW; 6DD33455B1F1B1C0 CRC64;

Query Match 18.6%; Score 8; DB 12; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TKS 4  
 | |  
 DB 2 TKS 4

## RESULT 14

P70804 PRELIMINARY; PRT; 7 AA.  
 AC P70804;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE ALGG GENE (FRAGMENT).  
 GN ALGT.  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Azotobacter.  
 OX NCBI\_TaxID=354;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-E;  
 RX MEDLINE=96427318; PubMed=8830682;  
 RA Rehm B.H.A., Ertesvag H., Valla S.;  
 RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algg) is  
 RT part of an alg gene cluster physically organized in a manner similar  
 RT to that in Pseudomonas aeruginosa.";  
 RL J. Bacteriol. 178:5884-5889(1996).  
 DR EMBL; X87973; CAA61230.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 16.3%; Score 7; DB 2; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TKS 4  
 | |  
 DB 2 TKS 4

## RESULT 15

O50556 PRELIMINARY; PRT; 7 AA.  
 AC O50556;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE GLYA (FRAGMENT).  
 GN GLYA.  
 OS Actinobacillus actinomycetemcomitans  
 OS (Haemophilus actinomycetemcomitans).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Actinobacillus.  
 OX NCBI\_TaxID=714;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 33384;  
 RX MEDLINE=96355846; PubMed=8751884;

RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,  
 RA Kraig E.;  
 RT "cis Elements and trans factors are both important in strain-specific  
 RT regulation of the leukotoxin gene in Actinobacillus  
 RT actinomycetemcomitans.";  
 RL Infect. Immun. 64:3451-3460(1996).  
 DR EMBL; U51862; AAB88721.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 16.3%; Score 7; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 Y 6  
 | |  
 DB 6 Y 6

Search completed: January 17, 2001, 13:49:02  
 Job time: 106 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:43:41 ; Search time 11.98 Seconds  
(without alignments)  
10.492 Million cell updates/sec

Title: US-08-765-837-7  
Perfect score: 43  
Sequence: 1 ETRSWYF 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues  
Total number of hits satisfying chosen parameters: 30043

Minimum DB seq length: 0  
Maximum DB seq length: 7

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pep: \*  
2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pep: \*  
3: /cgn2.6/ptodata/2/1aa/6\_COMB.pep: \*  
4: /cgn2.6/ptodata/2/1aa/Pctus\_COMB.pep: \*  
5: /cgn2.6/ptodata/2/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	25	58.1	6	1	US-07-653-091A-8
2	25	58.1	6	1	US-08-485-068-8
3	25	58.1	6	1	US-08-481-598-8
4	25	58.1	6	2	US-08-483-353-8
5	25	58.1	6	2	US-08-473-702-8
6	24	55.8	4	1	US-08-014-979-112
7	24	55.8	4	1	US-08-014-979-113
8	24	55.8	4	1	US-08-014-979-114
9	24	55.8	4	1	US-08-014-979-115
10	24	55.8	6	1	US-08-487-006-16
11	24	55.8	6	2	US-08-488-659A-16
12	23	53.5	4	3	US-09-260-190-6
13	23	53.5	5	3	US-09-260-190-4
14	23	53.5	6	1	US-08-487-006-108
15	23	53.5	6	2	US-08-488-659A-108
16	22	51.2	4	3	US-08-996-338-15
17	22	51.2	5	1	US-08-665-220-70
18	22	51.2	5	2	US-08-618-408B-63
19	22	51.2	5	5	5217869-89
20	22	51.2	6	2	US-08-687-956A-15
21	21	48.8	5	3	US-08-981-122-39
22	21	48.8	6	1	US-08-375-911A-20
23	21	48.8	6	1	US-08-487-006-20
24	21	48.8	6	2	US-08-488-659A-20
25	21	48.8	7	1	US-07-973-235A-30
26	21	48.8	7	2	US-08-462-720-30
27	20	46.5	5	3	US-08-981-122-47
28	20	46.5	5	3	US-08-981-122-54

29	20	46.5	6	1	US-08-321-625-28	Sequence 28, Appl
30	20	46.5	6	1	US-08-321-625-29	Sequence 29, Appl
31	20	46.5	6	1	US-08-321-625-78	Sequence 78, Appl
32	20	46.5	6	1	US-08-487-006-94	Sequence 94, Appl
33	20	46.5	6	1	US-08-433-318A-2	Sequence 2, Appl
34	20	46.5	6	1	US-08-433-318A-6	Sequence 6, Appl
35	20	46.5	6	1	US-08-467-607-12	Sequence 12, Appl
36	20	46.5	6	2	US-08-469-362-12	Sequence 12, Appl
37	20	46.5	6	2	US-08-850-392-12	Sequence 12, Appl
38	20	46.5	6	2	US-08-922-048-2	Sequence 2, Appl
39	20	46.5	6	2	US-08-922-048-6	Sequence 6, Appl
40	20	46.5	6	2	US-08-488-659A-94	Sequence 94, Appl
41	20	46.5	6	4	PCT-US96-06270-2	Sequence 2, Appl
42	20	46.5	6	4	PCT-US96-06270-6	Sequence 6, Appl
43	20	46.5	7	1	US-08-321-625-4	Sequence 4, Appl
44	20	46.5	7	1	US-08-321-625-26	Sequence 26, Appl
45	20	46.5	7	1	US-08-321-625-70	Sequence 70, Appl

ALIGNMENTS

RESULT 1  
US-07-653-091A-8  
; Sequence 8, Application US/07653091A  
; Patent No. 5614366  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.,  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: HTLV-II AND HTLV-II Peptide Antigens and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Swiss  
; STREET: P.O. Box 60850  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/653,091A  
; FILING DATE: 08-FEB-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/366,313  
; FILING DATE: 13-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/948,270  
; FILING DATE: 31-DEC-1986  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dehlinger, Peter J.  
; REGISTRATION NUMBER: 28,006  
; REFERENCE/DOCKET NUMBER: 4600-032.30  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 323-8302  
; TELEFAX: (415) 323-8306  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: HTLV-II  
US-07-653-091A-8

Query Match 58.1%; Score 25; DB 1; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SWYF 7  
|||:  
Db 2 SWY 5

RESULT 2  
US-08-485-068-8  
; Sequence 8, Application US/08485068  
; Patent No. 5763572  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: HTLV-I AND HTLV-II Peptide Antigens and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genelabs Technologies, Inc., Legal Dept.  
; STREET: 505 Penobscot Drive  
; CITY: Redwood City  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94063  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,068  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/653,091  
; FILING DATE: 08-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/366,313  
; FILING DATE: 13-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/948,270  
; FILING DATE: 31-DEC-1986  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, Allan A.  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: G4CIP2D1  
; TELEPHONE: (415) 369-9500  
; TELEFAX: (415) 368-0709  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: HTLV-II  
US-08-485-068-8

Query Match 58.1%; Score 25; DB 1; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SWYF 7  
|||:  
Db 2 SWY 5

RESULT 2  
US-08-485-068-8  
; Sequence 8, Application US/08485068  
; Patent No. 5763572  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: HTLV-I AND HTLV-II Peptide Antigens and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genelabs Technologies, Inc., Legal Dept.  
; STREET: 505 Penobscot Drive  
; CITY: Redwood City  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94063  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,068  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/653,091  
; FILING DATE: 08-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/366,313  
; FILING DATE: 13-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/948,270  
; FILING DATE: 31-DEC-1986  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, Allan A.  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: G4CIP2D1  
; TELEPHONE: (415) 369-9500  
; TELEFAX: (415) 368-0709  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: HTLV-II  
US-08-485-068-8

Db 2 SWY 5

RESULT 3  
US-08-481-598-8  
; Sequence 8, Application US/08481598  
; Patent No. 5814441  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: HTLV-I AND HTLV-II Peptide Antigens and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genelabs Technologies, Inc., Legal Dept.  
; STREET: 505 Penobscot Drive  
; CITY: Redwood City  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94063  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,598  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/653,091  
; FILING DATE: 08-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/366,313  
; FILING DATE: 13-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/948,270  
; FILING DATE: 31-DEC-1986  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, Allan A.  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: G4CIP2D1  
; TELEPHONE: (415) 369-9500  
; TELEFAX: (415) 368-0709  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: HTLV-II  
US-08-481-598-8

Query Match 58.1%; Score 25; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SWYF 7  
|||:  
Db 2 SWY 5

RESULT 4  
US-08-483-353-8  
; Sequence 8, Application US/08483353  
; Patent No. 5871933  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: HTLV-I AND HTLV-II Peptide Antigens and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genelabs Technologies, Inc., Legal Dept.  
; STREET: 505 Penobscot Drive  
; CITY: Redwood City  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94063  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,353  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/653,091  
; FILING DATE: 08-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/366,313  
; FILING DATE: 13-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/948,270  
; FILING DATE: 31-DEC-1986  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, Allan A.  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: G4CIP2D1  
; TELEPHONE: (415) 369-9500  
; TELEFAX: (415) 368-0709  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; ORGANISM: HTLV-II  
US-08-483-353-8

APPLICANT: Reyes, Gregory R.  
APPLICANT: Hadlock, Kenneth G.  
TITLE OF INVENTION: HTLV-II AND HTLV-II Peptide Antigens and  
METHODS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc., Legal Dept.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483.353  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/653,091  
FILING DATE: 08-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/366,313  
FILING DATE: 13-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/948,270  
FILING DATE: 31-DEC-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, Allan A.  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: G4C1P2D1  
TELEPHONE: (415) 369-9500  
TELEFAX: (415) 368-0709  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: HTLV-II  
US-08-483-353-8

Query Match 58.1%; Score 25; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SWVF 7  
Db 2 SWY 5

RESULT 5  
US-08-473-702-8  
Sequence 8, Application US/08473702  
Patent No. 5928861  
GENERAL INFORMATION:  
APPLICANT: Reyes, Gregory R.  
APPLICANT: Hadlock, Kenneth G.  
TITLE OF INVENTION: HTLV-II AND HTLV-II Peptide Antigens and  
METHODS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc., Legal Dept.  
STREET: 505 Penobscot Drive

CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,702  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/653,091  
FILING DATE: 08-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/366,313  
FILING DATE: 13-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/948,270  
FILING DATE: 31-DEC-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, Allan A.  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: G4C1P2D1  
TELEPHONE: (415) 369-9500  
TELEFAX: (415) 368-0709  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: HTLV-II  
US-08-473-702-8

Query Match 58.1%; Score 25; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SWVF 7  
Db 2 SWY 5

RESULT 6  
US-08-014-979-112  
Sequence 112, Application US/08014979  
Patent No. 5510240  
GENERAL INFORMATION:  
APPLICANT: Lam, Kit S. et al.  
TITLE OF INVENTION: Random Bio-Oligomer Library, A Method of  
Synthesis Thereof, and a Method of Use Thereof  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/014,979  
;; FILING DATE: 19930208  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Misrock, S. Leslie  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 7156-041  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212 790-9090  
;; TELEFAX: 212 869-8864/9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 112:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 1  
;; OTHER INFORMATION: /label= W  
;; OTHER INFORMATION: /note= "9-fluorenylmethoxycarbonyl (Fmoc)  
;; OTHER INFORMATION: bound "  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 2  
;; OTHER INFORMATION: /label= Y  
;; OTHER INFORMATION: /note= "(O-t-Bu) bound"  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 3  
;; OTHER INFORMATION: /label= F  
;; OTHER INFORMATION: /note= "linked with ONB (an organic molecule)"  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 4  
;; OTHER INFORMATION: /label= Xaa  
;; OTHER INFORMATION: /note= "Xaa is beta-Ala-ACA-EDA-pepSyn K"  
US-08-014-979-112

Query Match 55.8%; Score 24; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7  
   | | |  
Db 1 WYF 3

RESULT 7  
US-08-014-979-113  
; Sequence 113, Application US/08014979  
; Patent No. 5510240  
; GENERAL INFORMATION:  
; APPLICANT: Lam, Kit S. et al.  
; TITLE OF INVENTION: Random Bio-Oligomer Library, A Method of  
; OPERATION OF INVENTION: Synthesis Thereof, and a Method of Use Thereof  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/014,979  
;; FILING DATE: 19930208  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Misrock, S. Leslie  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 7156-041  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212 790-9090  
;; TELEFAX: 212 869-8864/9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 113:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 3  
;; OTHER INFORMATION: /label= F  
;; OTHER INFORMATION: /note= "linked with ONB (an organic molecule)"  
;; FEATURE:  
;; NAME/KEY: Modified-site  
;; LOCATION: 4  
;; OTHER INFORMATION: /label= Xaa  
;; OTHER INFORMATION: /note= "Xaa is beta-Ala-ACA-EDA-pepSyn K"  
US-08-014-979-113

Query Match 55.8%; Score 24; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WYF 7  
   | | |  
Db 1 WYF 3

RESULT 8  
US-08-014-979-114  
; Sequence 114, Application US/08014979  
; Patent No. 5510240  
; GENERAL INFORMATION:  
; APPLICANT: Lam, Kit S. et al.  
; TITLE OF INVENTION: Random Bio-Oligomer Library, A Method of  
; OPERATION OF INVENTION: Synthesis Thereof, and a Method of Use Thereof  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,979  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7156-041  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741

TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1  
; OTHER INFORMATION: /label= W  
; OTHER INFORMATION: /note= "Fmoc bound"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 2  
; OTHER INFORMATION: /label= Y  
; OTHER INFORMATION: /note= "(O-t-Bu) bound"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /label= F  
; OTHER INFORMATION: /note= "linked with ONB (an organic molecule)"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 4  
; OTHER INFORMATION: /label= Xaa  
; OTHER INFORMATION: /note= "-beta-Ala-ACA-4-MBHA"  
US-08-014-979-114

Query Match 55.8%; Score 24; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WYF 7  
|||  
Db 1 WYF 3

RESULT 9  
US-08-014-979-115  
; Sequence 115, Application US/08014979  
; Patent No. 5510240  
; GENERAL INFORMATION:  
; APPLICANT: Lam, Kit S. et al.  
; TITLE OF INVENTION: Random Bio-Oligomer Library, A Method of  
; TITLE OF INVENTION: Synthesis Thereof, and a Method of Use Thereof  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,979  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7156-041  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 115:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 3  
; OTHER INFORMATION: /label= F  
; OTHER INFORMATION: /note= "linked with ONB (an organic molecule)"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 4  
; OTHER INFORMATION: /label= Xaa  
; OTHER INFORMATION: /note= "-beta-Ala-ACA-4-MBHA"  
US-08-014-979-115

Query Match 55.8%; Score 24; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WYF 7  
|||  
Db 1 WYF 3

RESULT 10  
US-08-487-006-16  
; Sequence 16, Application US/08487006  
; Patent No. 5641861  
; GENERAL INFORMATION:  
; APPLICANT: Dooley, Colette T.  
; APPLICANT: Houghten, Richard A.  
; TITLE OF INVENTION: NO. 5641861e1 Mu Opioid Receptor Ligands:  
; TITLE OF INVENTION: Agonists and Antagonists  
; NUMBER OF SEQUENCES: 222  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,006  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-Tp 1706  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1  
; OTHER INFORMATION: /note= "Amino acid is acetylated at  
; OTHER INFORMATION: the N-terminal."  
; FEATURE:

; NAME/KEY: Peptide  
; LOCATION: 6  
; OTHER INFORMATION: /note= "Amino acid is amidated at  
; the C-terminal."  
US-08-487-006-16

Query Match 55.8%; Score 24; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 5 WYF 7  
|||  
Db 4 WYF 6

RESULT 11  
US-08-488-659A-16  
; Sequence 16, Application US/08488659A  
; Patent No. 5919897  
; GENERAL INFORMATION:  
; APPLICANT: Dooley, Colette T.  
; APPLICANT: Houghten, Richard A.  
; TITLE OF INVENTION: Mu Opioid Receptor Ligands:  
; Agonists and Antagonists  
; NUMBER OF SEQUENCES: 223  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,659A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-TP 1705  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1  
; OTHER INFORMATION: /note= "Amino acid is acetylated at  
; the N-terminal."  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 6  
; OTHER INFORMATION: /note= "Amino acid is amidated at  
; the C-terminal."  
US-08-488-659A-16

Query Match 55.8%; Score 24; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 5 WYF 7  
|||

Db 4 WYF 6  
Query Match 53.5%; Score 23; DB 3; Length 4;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0;

Qy 3 KSWY 6  
| | |  
Db 1 KEWY 4

RESULT 12  
US-09-260-190-6  
; Sequence 6, Application US/09260190  
; Patent No. 6096713  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; APPLICANT: Blasecki, John W.  
; TITLE OF INVENTION: Pharmaceutical Angiostatic Dipeptide  
; Compositions and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/260,190  
; FILING DATE: 01-MAR-1999  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/538,701  
; FILING DATE: 03-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/614,764  
; FILING DATE: 13-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Powers, Margaret A.  
; REGISTRATION NUMBER: 39,804  
; REFERENCE/DOCKET NUMBER: 015542-002320US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-260-190-6

Query Match 53.5%; Score 23; DB 3; Length 4;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0;

Qy 3 KSWY 6  
| | |  
Db 1 KEWY 4

RESULT 13  
US-09-260-190-4  
; Sequence 4, Application US/09260190  
; Patent No. 6096713  
; GENERAL INFORMATION:  
; APPLICANT: Green, Lawrence R.  
; APPLICANT: Blasecki, John W.  
; TITLE OF INVENTION: Pharmaceutical Angiostatic Dipeptide  
; Compositions and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP



STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/260,190  
FILING DATE: 01-MAR-1999  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/538,701  
FILING DATE: 03-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/614,764  
FILING DATE: 13-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Powers, Margaret A.  
REGISTRATION NUMBER: 39,804  
REFERENCE/DOCKET NUMBER: 015542-002320US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-260-190-4

Query Match 53.5%: Score 23; DB 3; Length 5;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KSWY 6  
Db 2 KSWY 5

RESULT 14  
US-08-487-006-108  
Sequence 108, Application US/08487006  
Patent No. 5641861  
GENERAL INFORMATION:  
APPLICANT: Dooley, Colette T.  
APPLICANT: Houghten, Richard A.  
TITLE OF INVENTION: No. 5641861 Mu Opioid Receptor Ligands:  
TITLE OF INVENTION: Agonists and Antagonists  
NUMBER OF SEQUENCES: 222  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,006  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-TP 1706  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 108:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1-6  
OTHER INFORMATION: /note= "All amino acids are D-amino  
OTHER INFORMATION: acids."  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 6  
OTHER INFORMATION: /note= "Amino acid is amidated at  
OTHER INFORMATION: the C-terminal."  
US-08-487-006-108

Query Match 53.5%: Score 23; DB 1; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SWYF 7  
Db 3 SWWF 6

RESULT 15  
US-08-488-659A-108  
Sequence 108, Application US/08488659A  
Patent No. 5919897  
GENERAL INFORMATION:  
APPLICANT: Dooley, Colette T.  
APPLICANT: Houghten, Richard A.  
TITLE OF INVENTION: Mu Opioid Receptor Ligands:  
TITLE OF INVENTION: Agonists and Antagonists  
NUMBER OF SEQUENCES: 223  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,659A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-TP 1705  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 108:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Peptide

LOCATION: 1..6  
OTHER INFORMATION: /note= "All amino acids are D-amino  
OTHER INFORMATION: acids."  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 6  
OTHER INFORMATION: /note= "Amino acid is amidated at  
OTHER INFORMATION: the C-terminal."  
US-08-488-659A-108

Query Match 53.5%; Score 23; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 SWVF 7  
||:|  
Db 3 SWWF 6

Search completed: January 17, 2001, 13:47:42  
Job time: 241 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:42:21 ; Search time 43.61 seconds  
(without alignments)  
7.057 Million cell updates/sec

Title: US-08-765-837-6

Perfect score: 49

Sequence: 1 MAPTKDEFD 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 61695

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_36:\*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	9	17 R90950	Factor VIII antigen
2	24	49.0	5	21 Y61841	Cadherin-8 cell ad
3	24	49.0	5	21 Y62453	Cadherin-15 cell a
4	24	49.0	8	16 R73400	Human TSH receptor
5	24	49.0	9	12 R10267	Human ventricular
6	24	49.0	9	17 W49525	Human leucocyte an
7	23	46.9	8	14 R43246	Cytotoxic T lympho
8	23	46.9	8	19 W54664	Peptide from P.fal
9	23	46.9	8	19 W48987	Immunoreactive pe
10	23	46.9	8	20 Y09316	Immunogenic peptid
11	23	46.9	9	18 W43880	Specific human leu
12	22	44.9	5	21 Y62517	Cadherin-15 cell a

13	22	44.9	5	21 Y62867	PB-cadherin cell a
14	22	44.9	8	12 R12310	ID2 plasmodium sp
15	22	44.9	8	14 R43245	Cytotoxic T lympho
16	22	44.9	8	16 R21394	Plasmodium activa
17	22	44.9	8	16 R78840	P. falciparum circ
18	22	44.9	8	17 R98613	Borna disease virus
19	22	44.9	8	19 W54663	Peptide from P.fal
20	22	44.9	8	20 Y03662	Amino acid sequen
21	22	44.9	9	17 W49542	Human leucocyte an
22	22	44.9	9	20 W97963	Human str. optocoma
23	22	44.9	9	21 Y76701	SCP-1 gene binding
24	21	42.9	5	21 Y73631	OB-cadherin cell a
25	21	42.9	5	21 Y60717	OB-cadherin cell a
26	21	42.9	5	21 Y61148	Cadherin-6 cell ad
27	21	42.9	5	21 Y62852	PB-cadherin cell a
28	21	42.9	7	9 P82888	N-terminal of hcr
29	21	42.9	7	14 R36869	Insulin-like grow
30	21	42.9	7	14 R43613	Peptide derived fr
31	21	42.9	7	16 R79640	Endocarditis spec
32	21	42.9	8	14 R36899	Insulin-like grow
33	21	42.9	8	14 R43460	Ro/SSA epitope 23
34	21	42.9	8	14 R43643	Peptide derived fr
35	21	42.9	8	15 R58438	Partial peptide 37
36	21	42.9	8	17 R95061	Multidomain prote
37	21	42.9	9	14 R36870	Insulin-like grow
38	21	42.9	9	14 R36895	Insulin-like grow
39	21	42.9	9	14 R36915	Insulin-like grow
40	21	42.9	9	14 R36923	Insulin-like grow
41	21	42.9	9	14 R43590	Peptide derived fr
42	21	42.9	9	14 R43614	Peptide derived fr
43	21	42.9	9	14 R43639	Peptide derived fr
44	21	42.9	9	15 R57538	Peptide (38) inhib
45	21	42.9	9	15 R57568	Peptide (68) inhib

#### ALIGNMENTS

##### RESULT 1

R90950  
ID R90950 standard; peptide; 9 AA.  
XX  
AC R90950;  
XX  
XX  
DT 09-SEP-1996 (first entry)  
XX  
DE Factor VIII antigenic peptide corresp. to residues Met1823-Asp1831.  
XX  
KW Factor VIII; modification; inhibitor activity; binding; antibody;  
KW von Willebrand factor; immune disorder.  
XX  
OS Synthetic.  
PN  
PN WO9602572-A2.  
XX  
PD 01-FEB-1996.  
XX  
PF 14-JUL-1995; 95WO-BE00068.  
XX  
PR 14-JUL-1994; 94BE-0000666.  
XX  
PR (CROI-) CROIX ROUGE BELGIOUE.  
XX  
PI Di Giambattista M, Laub R;  
XX  
DR WPI; 1996-1058A1/11.  
XX  
PT Factor VIII antigenic polypeptide fragments and epitope(s) - also  
PT inhibitors of factor VIII and anti-inhibitors, useful for e.g.  
PT preventing and treating immune disorders involving inhibition of  
PT factor VIII binding  
XX  
PS Claim 8; Page 31; 45pp; French.

XX Peptides R9045-64 are derived from the factor VIII protein, esp. from a  
 CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,  
 CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII  
 CC and derived peptides can be used to prevent the activity of inhibitors of  
 CC factor VIII binding to von Willebrand factor, esp. antibodies, thus  
 CC preventing or treating immune disorders.

XX Sequence 9 AA;

Query Match 100.0%; Score 49; DB 17; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPTKDEED 9  
 Db 1 maptkdefd 9

RESULT 2

ID Y61841 standard; Peptide; 5 AA.

XX AC Y61841;

XX DT 02-MAR-2000 (first entry)

XX DE Cadherin-8 cell adhesion recognition cyclic peptide SEQ ID NO:1679.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Modified-site 1..5  
 FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"

XX WO9957149-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00363.

XX 05-MAY-1998; 98US-0073040.

XX 06-NOV-1998; 98US-0187859.

XX 20-JAN-1999; 99US-0234395.

XX 08-MAR-1999; 99US-0264516.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-038791/03.

XX New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease

XX • Claim 42; Page 177; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)

CC comprising peptides which comprise a nonclassical cadherin cell adhesion

CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Y33183 to Y33186 represent sequences used in the  
 CC exemplification of the present invention.

XX Sequence 5 AA;

Query Match 49.0%; Score 24; DB 21; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KDEFD 9  
 Db 1 keefd 5

RESULT 3

ID Y62453 standard; Peptide; 5 AA.

XX AC Y62453;

XX DT 02-MAR-2000 (first entry)

XX DE Cadherin-15 cell adhesion recognition cyclic peptide SEQ ID NO:1118.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Modified-site 1..5  
 FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"

XX WO9957149-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00363.

XX 05-MAY-1998; 98US-0073040.

XX 06-NOV-1998; 98US-0187859.

XX 20-JAN-1999; 99US-0234395.

XX 08-MAR-1999; 99US-0264516.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 XX  
 XX New cadherin modulating agents, used for modulating nonclassical  
 XX cadherin-mediated functions for treating e.g. cancers, obesity,  
 XX rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 XX disease -  
 XX  
 XX Claim 60; Page 188; 252pp; English.  
 XX  
 XX The present invention describes cadherin modulating agents (MA)  
 XX comprising peptides which comprise a nonclassical cadherin cell adhesion  
 XX recognition (CAR) sequence. The MAs can be used for modulating  
 XX nonclassical cadherin-mediated functions. They can be used for e.g.  
 XX inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 XX mammal, enhancing delivery of a drug to a tumour in a mammal, treating  
 XX a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 XX angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 XX expressing cell, preventing or treating obesity in a mammal, stimulating  
 XX blood vessel regression in a mammal, enhancing drug delivery to the  
 XX central nervous system, treating a demyelinating neurological disease,  
 XX increasing vasopermeability in a mammal, enhancing adhesion of  
 XX nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 XX a mammal, or preventing pregnancy in a mammal. They can also be used for  
 XX e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 XX or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 XX mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 XX age-related macular degeneration, multiple sclerosis and diabetes. The  
 XX products can also be used for detection and diagnosis and in  
 XX bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 XX and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 XX exemplification of the present invention.  
 XX  
 XX Sequence 5 AA;

Query Match 49.0%; Score 24; DB 21; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KDEFD 9  
 II:II  
 Db 1 kdkfd 5

RESULT 4  
 ID R73400 standard; Peptide; 8 AA.  
 XX AC R73400;  
 XX  
 XX DT 08-DEC-1995 (first entry)  
 XX  
 XX DE Human TSH receptor (residues 399-406).  
 XX  
 XX KW thyroid stimulating hormone receptor; TSH; human; Homo sapiens;  
 XX KW antibody; affinity; detection.  
 XX OS Synthetic.  
 XX PN JP07089991-A.  
 XX  
 XX PD 04-APR-1995.  
 XX  
 XX PF 28-SEP-1993; 93JP-0240853.  
 XX  
 XX PR 28-SEP-1993; 93JP-0240853.  
 XX  
 XX PA (MITSUBISHI PETROCHEMICAL CO LTD.  
 XX

DR WPI; 1995-167251/22.  
 XX  
 XX Novel polypeptide(s) having affinity for the human TSH receptor  
 XX antibody - used in detection of the TSH antibody.  
 XX  
 XX Example 1; Page 31; 54pp; Japanese.  
 XX  
 XX Peptides with affinity to human TSH (thyroid stimulating hormone);  
 XX receptor antibody are used for detection of the antibody. (See also  
 XX R73201-592).  
 XX  
 XX Sequence 8 AA;

Query Match 49.0%; Score 24; DB 16; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PTKDEFD 9  
 I III:  
 Db 2 pksdefn 8

RESULT 5  
 ID R10267 standard; Protein; 9 AA.  
 XX AC R10267;  
 XX  
 XX DT 21-MAR-1991 (first entry)  
 XX  
 XX DE Human ventricular myosin light chain-1 antigen (11).  
 XX  
 XX KW Human ventricular myosin light chain-1; antigen; monoclonal antibody;  
 XX immunosay; myocardial infarction; diagnosis; cardiomyopathy.  
 XX  
 XX OS Synthetic.  
 XX  
 XX PN W09015993-A.  
 XX  
 XX PD 27-DEC-1990.  
 XX  
 XX PF 14-JUN-1990; 90WO-US03434.  
 XX  
 XX PR 14-JUN-1989; 89US-0366913.  
 XX  
 XX PA (GEHO-) GEN HOSPITAL CORP.  
 XX  
 XX PI Khaw BA, Nicol PD, Matsueda GR;  
 XX  
 XX WPI; 1991-022323/03.  
 XX  
 XX Monoclonal antibody to human ventricular myosin light chain-1 -  
 XX used in immunoassays to diagnose myocardial infarction or other  
 XX cardio-myopathies  
 XX  
 XX Claim 5; Page 16; 34pp; English.  
 XX  
 XX The peptide is prep'd by solid phase synthesis and then coupled to a  
 XX carrier, esp. by a thioether linkage to a cysteine residue added to  
 XX the C-terminal or A-terminal end of the peptides, and used as  
 XX immunogens in animals to produce immune lymphocytes  
 XX  
 XX A pure monoclonal antibody is capable of forming an immune complex  
 XX with hVM IC-1 and has no capability of forming an immune complex  
 XX human fast skeletal myosin IC-1.  
 XX  
 XX hVM IC-1 is diagnostic of myocardial infarction or other  
 XX cardiomyopathies.  
 XX  
 XX See also R10266-73.  
 XX  
 XX Sequence 9 AA;

Query Match 49.0%; Score 24; DB 12; Length 9;

```

Best Local Similarity 71.4%; Pred. No. 2.1e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAPTKDE 7
   1 1 1 1 1
Db 1 mapkpe 7

RESULT 6
W49525
ID W49525 standard; peptide; 9 AA.
AC W49525;
XX
XX
DT 05-JUN-1998 (first entry)
XX
DE Human leucocyte antigen DQ4 binding peptide #416.
XX
XX Human leucocyte antigen; HLA-DQ4; combinatorial library;
KW autoimmune disease; chronic articular rheumatism.
XX
OS Synthetic.
XX
XX JP08151396-A.
XX
PD 11-JUN-1996.
XX
XX 28-NOV-1994; 94JP-0292657.
XX
PR 28-NOV-1994; 94JP-0292657.
XX
PA (TEIJ ) TEIJIN LTD.
XX
DR WPI; 1996-329479/33.
XX
XX HLA-binding oligopeptide and an immuno:regulator contg it - used in
PT the treatment of auto:immune disease
XX
PS Claim 4; Page 47; 61pp; Japanese.
XX
CC This peptide is an example of a peptide which binds to a human leucocyte
CC antigen HLA-DQ4 molecule. The peptide was isolated from a phagemid
CC combinatorial library comprising the sequence V05953, by screening with
CC an HLA-DQ4 molecule. The peptide is used for the treatment of autoimmune
CC disease, or especially for treatment of viral diseases.
XX
SQ Sequence 9 AA;

Query Match 49.0%; Score 24; DB 17; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.1e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 APTKDE 7
   1 1 1 1 1
Db 3 aptkek 8

RESULT 7
R43246
ID R43246 standard; peptide; 8 AA.
XX
AC R43246;
XX
XX
DT 04-MAY-1994 (first entry)
XX
DE Cytotoxic T lymphocyte recognition/induction peptide.
XX
XX CTL; vaccine; malaria; specific antigen-derived.
XX
OS Synthetic.
XX
FN W09330103-A.

us-08-765-837-6.closed.rag

XX 14-OCT-1993.
XX
XX 05-APR-1993; 93WO-GB00711.
XX
PR 03-APR-1992; 92GB-0008068.
PR 20-AUG-1992; 92GB-0017704.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
XX Elvin J, Gotch FM, Hill AV, McMichael AJ, Whittle HC;
PI WPI; 1993-336833/42.
XX
XX Peptide(s) recognising or inducing cytotoxic T lymphocytes -
PT useful in vaccines against malaria or HIV-2, derived from
PT specific antigen and human leukocyte antigen contg. class I
PT restricted epitope
XX
XX Claim 6; Page 30; 35pp; English.
XX
XX The sequence is that of peptide cp29 which is recognised by, or can
CC induce, cytotoxic T lymphocytes. It may be useful in vaccines against
CC malaria.
XX
SQ Sequence 8 AA;

Query Match 46.9%; Score 23; DB 14; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.1e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 TKDFED 9
   1 1 1 1 1
Db 2 skdeid 7

RESULT 8
W54664
ID W54664 standard; peptide; 8 AA.
XX
AC W54664;
XX
DT 25-SEP-1998 (first entry)
XX
DE Peptide from P.falciparum CSP 368-375 b.
XX
XX Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;
KW vaccine; treatment.
XX
OS Synthetic.
XX
PN W09813378-A1.
XX
PD 02-APR-1998.
XX
XX 25-SEP-1997; 97WO-NL00536.
XX
PR 26-SEP-1996; 96EP-0202701.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
XX
XX Drijfhout JW, Koning F;
XX WPI; 1998-230631/20.
XX
XX Increasing uptake and presentation of antigen(s) - by adding mannose
PT residue(s) to antigen for increasing T cell response, useful in,
PT e.g. vaccines against viral infection(s)
XX
PS Disclosure; Page 29; 47pp; English.
XX
XX The peptides W54559-W54809 are examples of peptides to which at least 1

```

CC (preferably 2) mannose can be attached to increase their uptake as  
 CC antigens by antigen-presenting cells. Uptake of agonist mannosylated  
 CC peptides will increase the T cell response, whereas uptake of antagonist  
 CC peptides blocks the T cell response. Blocking binding of immunogenic  
 CC autoantigens can be used in treatment of type I diabetes, rheumatoid  
 CC arthritis, graft rejection etc., also to induce T-cell non-  
 CC responsiveness. Vaccines containing mannosylated antigen are used to  
 CC prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths  
 CC and parasites.

XX Sequence 8 AA;

Query Match 46.9%; Score 23; DB 19; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TKDEFD 9  
 Db :||| 1  
 2 skdel 7

RESULT 9  
 W48987  
 ID W48987 standard; peptide; 8 AA.  
 AC W48987;  
 XX  
 XX 29-SEP-1998 (first entry)  
 DT  
 XX  
 XX Immuno-reactive peptide CTL epitope 4 of human cytomegalovirus.  
 DE  
 XX  
 KW Immunogenic epitope; human cytomegalovirus; HCMV; vaccine; CTL;  
 KW CD8+; cytotoxic T-lymphocyte; MHC; major histocompatibility complex.  
 XX  
 XX Synthetic.  
 OS Human cytomegalovirus.  
 PN W09821233-A2.  
 XX  
 XX 22-MAY-1998.

XX 12-NOV-1997; 97WO-US20236.  
 PF  
 XX 14-OCT-1997; 97US-0950064.  
 PR  
 XX 12-NOV-1996; 96US-0747488.  
 PR  
 PA (CITY ) CITY OF HOPE.  
 XX  
 PI Diamond DJ, York J;  
 XX  
 DR WPI; 1998-297862/26.  
 XX  
 XX Immunoreactive human cytomegalovirus epitopes - useful to vaccinate  
 PT against infection  
 PT  
 XX Claim 1; Page 48; 56pp; English.

XX The present peptide is an immunogenic epitope which is recognised  
 CC by the CD8+ class I major histocompatibility complex (MHC) restricted  
 CC cytotoxic T-lymphocytes of patients harboring latent HCMV infection.  
 CC The peptide is claimed to be capable of activating cytotoxic T-cell  
 CC lymphocytes in the absence of active viral replication, and thus is  
 CC useful for eliciting a cellular immune response against HCMV by normal  
 CC and immunodeficient subjects. The immunological peptide can also be  
 CC formulated as a vaccine which is claimed to be useful for protecting  
 CC against HCMV infection, augmenting the immune system response to a HCMV  
 CC infection or protecting against reactivation of a latent HCMV infection.

XX Sequence 8 AA;

Query Match 46.9%; Score 23; DB 19; Length 8;

Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PTKD 6  
 Db :||| 1  
 2 ptkd 5

RESULT 10  
 Y09316  
 ID Y09316 standard; peptide; 8 AA.  
 AC Y09316;  
 XX  
 XX 08-JUL-1999 (first entry)  
 DT  
 XX  
 XX Immunogenic peptide cytotoxic T lymphocyte epitope SEQ ID NO:9 of HCMV.  
 DE  
 XX Human cytomegalovirus; HCMV; immunologically active peptide; vaccine;  
 KW immune response; cytotoxic T lymphocyte; CTL; immunostimulation;  
 KW infection; immunosuppression; bone marrow transplant; solid organ;  
 KW heart; AIDS.  
 XX  
 OS Human cytomegalovirus.  
 OS Synthetic.  
 XX  
 PN W09919349-A1.  
 XX  
 XX 22-APR-1999.  
 PD  
 XX 11-MAY-1998; 98WO-US09652.  
 PF  
 XX 10-FEB-1998; 98US-0021298.  
 PR  
 PR 14-OCT-1997; 97US-0950064.  
 XX  
 XX (CITY ) CITY OF HOPE.  
 PA  
 XX Diamond DJ, York J;  
 PI  
 XX WPI; 1999-277590/23.  
 DR

XX Immunogenic peptide cytotoxic T lymphocyte epitopes of human  
 PT cytomegalovirus  
 PT  
 XX Claim 1; Page 49; 64pp; English.  
 PS  
 XX The present invention describes immunologically active peptides (IAPs)  
 CC capable of eliciting a cellular immune response to human cytomegalovirus  
 CC (HCMV). The IAP can be used in a (cellular) vaccine to augment the  
 CC immune system response to HCMV, or to provide immunity against HCMV. The  
 CC IAP (cellular) vaccine can also protect an individual having a latent  
 CC HCMV infection from reactivation. An antigen presenting cell can be used  
 CC to determine the presence or absence of HCMV-infected T lymphocytes. The  
 CC viral vector containing IAP encoding DNA can also be used to provide  
 CC immunity against HCMV. The IAP can be used to prepare HCMV-reactive  
 CC human cytotoxic T lymphocytes. The antigen presenting cells primed with  
 CC the IAPs can be used as diagnostic reagents to detect immunostimulation  
 CC by HCMV. They can also detect active HCMV infection or exposure to HCMV.  
 CC HCMV can cause opportunistic infections resulting in a variety of  
 CC complications in, e.g. immunosuppressed patients. The IAP vaccines  
 CC impart immunity to bone marrow transplant recipients, solid organ  
 CC recipients, heart patients, AIDS patients or women of child-bearing  
 CC years, without the need for ex vivo expansion of HCMV-specific cytotoxic  
 CC T lymphocytes (which requires sophisticated laboratory setup and is  
 CC highly labor-intensive and costly).

Sequence 8 AA;

Query Match 46.9%; Score 23; DB 20; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTKD 6  
 ID 1111  
 Db 2 ptkd 5

RESULT 11  
 W43880  
 ID W43880 standard; peptide; 9 AA.  
 AC  
 XX  
 XX  
 DT 20-APR-1998 (first entry)  
 XX  
 DE  
 XX  
 KW Human leukocyte antigen binding peptide #84.  
 KW Human leukocyte antigen; HLA; vaccine; bacterial infection; viral;  
 KW cytotoxic T-cell; CTL; immunogenic peptide; cancer.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN W09734617-A1.  
 PD 25-SEP-1997.  
 XX  
 PF 21-MAR-1997; 97WO-US04451.  
 XX  
 PR 20-MAR-1997; 97US-0821739.  
 PR 21-MAR-1996; 96US-0013833.  
 XX  
 XX (CYTE-) CYTEL CORP.  
 PA  
 XX  
 PI Celis E, Grey HM, Kubo RT, Sette A;  
 XX WPI; 1997-489250/45.  
 DR  
 XX  
 PT Specific human leukocyte antigen binding peptide - used in vaccines  
 PT for the treatment and prevention of e.g. bacterial or viral  
 PT infection and cancer  
 XX  
 XX Claim 19; Page 39; 49pp; English.  
 PS  
 CC The present sequence represents a specific example of an immunogenic  
 CC peptide which was used in a new method of inducing a cytotoxic T cell  
 CC (CTL) response against a preselected antigen in a patient. The method  
 CC comprises contacting CTLs from the patient with the immunogenic peptide  
 CC (containing defined motifs) which binds one of the four HLA MHC products  
 CC HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1, the peptide having a dissociation  
 CC constant (Kd) of less than 5 x 10<sup>-7</sup> M. Immunogens are viral, e.g. human  
 CC immunodeficiency virus type I (HIV-1), hepatitis B virus (HBV) and  
 CC hepatitis C virus (HCV) or cancer antigens and are used in vaccines for  
 CC the prevention and treatment of viral infection and cancer. The  
 CC immunogens may be administered to the patient as a nucleic acid encoding  
 CC the peptide (gene vaccine).  
 XX  
 SQ Sequence 9 AA;  
 Query Match 46.9%; Score 23; DB 18; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 APTKDEF 8  
 ID 1111  
 Db 3 agtkqef 9

RESULT 12  
 Y62517  
 ID Y62517 standard; Peptide; 5 AA.  
 AC  
 XX  
 XX Y62517;  
 XX

DT 02-MAR-2000 (first entry)  
 XX  
 DE  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 XX Synthetic.  
 OS  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1..5  
 FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"  
 XX  
 XX W09957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 XX 05-MAY-1999; 99WO-CA00363.  
 XX  
 PF 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
 PA  
 XX Blaschuk OW, Gour BJ, Byers S;  
 PI  
 XX WPI; 2000-038791/03.  
 DR  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease  
 XX  
 XX Claim 60; Page 188; 252pp; English.  
 PS  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z3183 to Z3186 represent sequences used in the  
 CC exemplification of the present invention.  
 XX  
 XX Sequence 5 AA;  
 SQ

Query Match 44.9%; Score 22; DB 21; Length 5;



Best Local Similarity 80.0%; Pred. No. 2.1e+05; Mismatches 1; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KDEFD 9  
1111

Db 1 kdel 5

RESULT 13  
Y62867

ID Y62867 standard; Peptide; 5 AA.

AC Y62867;

XX 02-MAR-2000 (first entry)

XX PB-cadherin cell adhesion recognition cyclic peptide SPQ ID NO:2423.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;

KW Inhibition; cadherin extracellular domain; cell adhesion recognition;

KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;

KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;

KW cadherin related neuronal receptor; Li-cadherin; protocadherin;

KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;

KW Rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;

KW neurological disease; cyclic.

XX Synthetic.

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Modified-site 1..5

FT /note= "the terminal residues are condensed with each

FT other to form a cyclic peptide"

XX

XX W09957149-A2.

XX 11-NOV-1999.

XX

XX 05-MAY-1999; 99WO-CA00363.

XX

XX 05-MAY-1998; 98US-0073040.

PR 06-NOV-1998; 98US-0187859.

PR 20-JAN-1999; 99US-0234395.

PR 08-MAR-1999; 99US-0264516.

XX

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX

XX Blaschuk OW, Gour BJ, Byers S;

XX

XX WPI; 2000-038791/03.

XX

XX New cadherin modulating agents, used for modulating nonclassical

PT cadherin-mediated functions for treating e.g. cancers, obesity,

PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological

PT disease

XX

XX Claim 72; Page 194; 252pp; English.

XX

XX The present invention describes cadherin modulating agents (MA)

CC comprising peptides which comprise a nonclassical cadherin cell adhesion

CC recognition (CAR) sequence. The MAs can be used for modulating

CC nonclassical cadherin-mediated functions. They can be used for e.g.

CC inhibiting adhesion of nonclassical-cadherin expressing cells in a

CC mammal, enhancing delivery of a drug through the skin of a mammal,

CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in

CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting

CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-

CC expressing cell, preventing or treating obesity in a mammal, stimulating

CC blood vessel regression in a mammal, enhancing drug delivery to the

CC central nervous system, treating a demyelinating neurological disease,

CC increasing vasopermeability in a mammal, enhancing adhesion of

CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in

CC

CC a mammal, or preventing pregnancy in a mammal. They can also be used for

CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing

CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a

CC mammal. They can also be used for treating e.g. psoriasis, arthritis,

CC age-related macular degeneration, multiple sclerosis and diabetes. The

CC products can also be used for detection and diagnosis and in

CC bioassays. Y60592 to Y64572 represent specifically claimed peptides,

CC and Y64573 to Y64643 and Y33183 to Y33186 represent sequences used in the

CC exemplification of the present invention.

XX

SQ Sequence 5 AA;

Query Match 44.9%; Score 22; DB 21; Length 5;

Best Local Similarity 80.0%; Pred. No. 2.1e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KDEFD 9  
1111

Db 1 kdel 5

RESULT 14  
R12310

ID R12310 standard; Protein; 8 AA.

XX

AC R12310;

XX

XX 29-AUG-1991 (first entry)

XX

XX ID2 plasmodium surface peptide.

DE

XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;

KW hybrid.

XX

XX Plasmodium falciparum.

OS

XX EP432965-A.

XX

XX 19-JUN-1991.

XX

XX 06-DEC-1990; 90EP-0313257.

PF

XX 08-DEC-1989; 89US-0447746.

PR

XX (SMIK ) SMITHKLINE BEECHAM.

PA (USSA ) US SEC OF THE ARMY.

PA (BIOM-) BIOMEDICAL RES INST.

XX

XX Gross MS, Gordon DM, Hollingdale MR;

PI

XX WPI; 1991-179771/25.

XX

XX Polypeptide comprising immunogenic determinants from P falciparum

PT - for vaccine against malaria infection in humans.

XX

XX Claim 2; Page 16; 18pp; English.

XX

XX The peptide is a sporozoite neutralising epitope from the 2nd flank-

CC ing region of the CS protein of plasmodium. It can be used in a

CC vaccine for protection against malaria. The peptide is pref. linked

CC to a 2nd peptide from the 1st flanking domain and the resulting

CC polypeptide fused to a carrier protein, e.g. tetanus toxoid, dip-

CC theria toxin or cholera B toxin. A preferred vaccine comprises 81

CC N-terminal AAs of the influenza virus nonstructural protein 1

CC (NS181), fused, via a synthetic linker, to a 1st flanking region of

CC the CS protein, which is itself fused to a 2nd flanking region of

CC the CS protein.

XX See also R12306-R12311 and R13175-R13179.

XX

SQ Sequence 8 AA;

Query Match 44.9%; Score 22; DB 12; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 KDEFD 9  
    | | | |  
Db 4 kdel 8

## RESULT 15

R43245  
ID R43245 standard; peptide; 8 AA.

XX AC R43245;

XX XX  
DT 04-MAY-1994 (first entry)

XX Cytotoxic T lymphocyte recognition/induction peptide.

XX CTL; vaccine; malaria; specific antigen-derived.

XX OS Synthetic.

XX PN W09320103-A.

XX PD 14-OCT-1993.

XX PF 05-APR-1993; 93WO-GB00711.

XX PR 03-APR-1992; 92GB-0008068.

XX PR 20-AUG-1992; 92GB-0017704.

XX PA (ISIS-) ISIS INNOVATION LTD.

XX PI Elvin J, Gotch FM, Hill AV, McMichael AJ, Whittle HC;

XX DR WPI; 1993-336833/42.

XX PT Peptide(s) recognising or inducing cytotoxic T lymphocytes -  
useful in vaccines against malaria or HIV-2, derived from  
specific antigen and human leukocyte antigen contg. class I  
restricted epitope

XX PS Claim 6; Page 30; 35pp; English.

XX CC The sequence is that of peptide cp26 which is recognised by, or can  
induce, cytotoxic T lymphocytes. It may be useful in vaccines against  
malaria.

XX SQ Sequence 8 AA;

Query Match 44.9%; Score 22; DB 14; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 KDEFD 9  
    | | | |  
Db 3 kdel 7

Search completed: January 17, 2001, 13:42:22  
Job time: 136 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:43:36 ; Search time 36.59 Seconds  
(without alignments)  
16.701 Million cell updates/sec

Title: us-08-765-837-6

Perfect score: 49

Sequence: 1 MAPTKDEED 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 787

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_66:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	44.9	9	S66635	alpha-2-macroglobu
2	22	44.9	9	S66636	alpha-2-macroglobu
3	18	36.7	4	D41654	hypothetical prote
4	18	36.7	9	PL0139	carbon-monoxide de
5	17	34.7	6	JH0784	neuropeptide TE-6
6	17	34.7	6	S15596	orf 3 rata 5'-regi
7	17	34.7	9	B41983	orf downstream to b
8	16	32.7	5	PS0324	ribulose-bisphosph
9	16	32.7	8	PT0530	T-cell receptor be
10	16	32.7	9	D48186	ATPase R1 subunit
11	16	32.7	9	C41978	callifMRamide 3 -
12	16	32.7	9	E41978	callifMRamide 5 -
13	16	32.7	9	B38740	Ig kappa chain C r
14	16	32.7	9	A42266	peptidylglycine mo
15	15	30.6	6	B60110	repetitive protein
16	15	30.6	8	S29272	tocopherol-binding
17	15	30.6	9	A44787	callifMRamide 10
18	15	30.6	9	B41978	callifMRamide 2 -
19	15	30.6	9	PC7078	unidentified 48.7K
20	15	30.6	9	PH0937	T-cell receptor be
21	14	28.6	7	B39040	calsequestrin, fas
22	14	28.6	9	D41978	callifMRamide 4 -
23	13	26.5	7	S57274	triacylglycerol li
24	13	26.5	8	S21273	cellulase (EC 3.2.
25	13	26.5	8	A61328	trypsin (EC 3.4.21
26	13	26.5	9	A41978	callifMRamide 1 -
27	12	24.5	3	I78890	tyrosine protein k
28	12	24.5	4	I40697	biotin A - Citroba
29	12	24.5	4	S53508	starvation-induced

30	12	24.5	5	1	H0ROHA	proctolin Americ
31	12	24.5	5	2	P00009	angiotensin-conver
32	12	24.5	5	2	A44692	fulicin - giant Af
33	12	24.5	5	2	A60411	proctolin - Atlant
34	12	24.5	5	2	PT0656	T-cell receptor be
35	12	24.5	6	2	I65546	MHC H2-L antigen -
36	12	24.5	6	2	PT0510	T-cell receptor be
37	12	24.5	6	2	PT0599	T-cell receptor be
38	12	24.5	7	2	S16365	opacity protein P.
39	12	24.5	7	2	B44787	callifMRamide 11
40	12	24.5	7	2	S68004	hucolin, 75K chain
41	12	24.5	8	2	A31570	angiotensin-conver
42	12	24.5	8	2	S53008	citrate synthase -
43	12	24.5	8	2	S11078	glucose-6-phosphat
44	12	24.5	8	2	B33099	158K exoantigen -
45	12	24.5	8	2	S71919	alcohol dehydrogen

ALIGNMENTS

RESULT 1

S66635

alpha-2-macroglobulin isoform 1 - bovine (fragment)

C:Species: Bos primigenius indicus (zebu cattle)

C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999

C:Accession: S66635

R:Dolmer, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Sot

FEBS Lett. 372, 93-95, 1995

A:Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain

A:Reference number: S66634; MUID:96032553

A:Accession: S66635

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <DOL>

Query Match 44.9%; Score 22; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 KDEF 8

Db 1 KDEF 4

RESULT 2

S66636

alpha-2-macroglobulin isoform 2 - bovine (fragment)

C:Species: Bos primigenius indicus (zebu cattle)

C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 17-Mar-1999

C:Accession: S66636

R:Dolmer, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Sot

FEBS Lett. 372, 93-95, 1995

A:Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain

A:Reference number: S66634; MUID:96032553

A:Accession: S66636

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <DOL>

Query Match 44.9%; Score 22; DB 2; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 PTKDE 7

Db 5 PKRDE 9

RESULT 3

D41654

hypothetical protein (sodC 5' region) - Haemophilus parainfluenzae (fragment)  
 C:Species: Haemophilus parainfluenzae  
 C:Date: 12-Jun-1992 #sequence\_revision 12-Jun-1992 #text\_change 24-Feb-1995  
 C:Accession: D41654  
 R:Kroll, J.S.; Langford, P.R.; Loynds, B.M.  
 J. Bacteriol. 173, 7449-7457, 1991  
 A:Title: Copper-zinc superoxide dismutase of Haemophilus influenzae and Haemophilus para  
 A:Reference number: A41654; MUID:92041655  
 A:Accession: D41654  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-4 <KRO>

Query Match 36.7%; Score 18; DB 2; Length 4;  
 Best Local Similarity 75.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 6 DEFD 9  
 1 11  
 1 DRFD 4

RESULT 4  
 PL0139  
 Carbon-monoxide dehydrogenase (EC 1.2.99.2) large chain - Pseudomonas carboxydoflava (fr  
 C:Species: Pseudomonas carboxydoflava  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 28-Apr-1993  
 C:Accession: PL0139  
 R:Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.  
 Arch. Microbiol. 152, 335-341, 1989  
 A:Title: Homology and distribution of CO dehydrogenase structural genes in carboxydotro  
 A:Reference number: PL0138; MUID:90055678  
 A:Accession: PL0139  
 A:Molecule type: protein  
 A:Residues: 1-9 <KRA>  
 A:Note: 2-Met is also found  
 C:Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large, me  
 C:Keywords: oxidoreductase

Query Match 36.7%; Score 18; DB 2; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 APTKD 6  
 1 1  
 3 APVQD 7

RESULT 5  
 JH0784  
 neuropeptide TE-6 - pig roundworm (fragment)  
 C:Species: Ascaris suum (pig roundworm)  
 C:Date: 10-Jun-1993 #sequence\_revision 10-Jun-1993 #text\_change 28-Apr-1995  
 C:Accession: JH0784  
 R:Smart, D.; Shaw, C.; Curry, W.J.; Johnston, C.F.; Thim, L.; Halton, D.W.; Buchanan, K.  
 Biochem. Biophys. Res. Commun. 187, 1323-1329, 1992  
 A:Title: The primary structure of TE-6: a novel neuropeptide from the nematode Ascaris s  
 A:Reference number: JH0784; MUID:93038603  
 A:Accession: JH0784  
 A:Molecule type: protein  
 A:Residues: 1-6 <SMA>  
 A:Experimental source: gonoduct  
 C:Keywords: neuropeptide

Query Match 34.7%; Score 17; DB 2; Length 6;  
 Best Local Similarity 50.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 KDFD 9  
 1 1 1

Db 1 TKOELE 6  
 RESULT 6  
 SI5596  
 orf 3 rara 5'-region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Jun-1999 #sequence\_revision 04-Jun-1999 #text\_change 28-Jun-1999  
 C:Accession: SI5596  
 R:Brand, N.J.; Petkovich, M.; Chambon, P.  
 Nucleic Acids Res. 18, 6799-6806, 1990  
 A:Title: Characterization of a functional promoter for the human retinoic acid recept  
 A:Reference number: SI5594; MUID:91088249  
 A:Accession: SI5596  
 A:Molecule type: DNA  
 A:Residues: 1-6 <BRA>  
 A:Cross-references: EMBL:X56058; NID:g35876  
 A:Note: This ORF from Fig. 2 is not annotated in GenBank entry HSRARA2, release 111.0  
 C:Comment: This sequence is not thought to be translated.  
 C:Genetics:  
 A:Gene: GDB:RARA  
 A:Cross-references: GDB:120337; OMIM:180240  
 A:Map position: 17q12-17q12

Query Match 34.7%; Score 17; DB 4; Length 6;  
 Best Local Similarity 75.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAPTK 4  
 1 11  
 1 MAPS 4

Db 1 MAPS 4

RESULT 7  
 B41983  
 orf downstream to bacterioferritin - Azotobacter vinelandii (fragment)  
 C:Species: Azotobacter vinelandii  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Oct-1999  
 C:Accession: B41983  
 R:Grossman, M.J.; Hinton, S.M.; Minak-Bernero, V.; Slaughter, C.; Stiefel, E.L.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 2419-2423, 1992  
 A:Title: Unification of the ferritin family of proteins.  
 A:Reference number: A41983; MUID:92196129  
 A:Accession: B41983  
 A>Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid; protein  
 A:Residues: 1-9 <GRO>  
 A:Cross-references: GB:M83692; NID:gl42297; PIDN:AAA22122.1; PID:gl42299  
 A:Note: sequence extracted from NCBI backbone (NCBIP:88442)

Query Match 34.7%; Score 17; DB 2; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPTK 5  
 1 11  
 1 MAPPR 5

Db 1 MAPPR 5

RESULT 8  
 PS0324  
 ribulose-bisphosphate carboxylase activase III - rice (strain Nihonbare) (fragment)  
 C:Species: Oryza sativa (rice)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 02-Jul-1998  
 C:Accession: PS0324  
 R:Tsuigita, A.  
 submitted to JIPID, April 1993  
 A:Reference number: PS0206  
 A:Accession: PS0324  
 A:Molecule type: protein  
 A:Residues: 1-5 <TSU>

A:Experimental source: leaf, chlorophyll

Query Match 32.7%; Score 16; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 APT 4  
|||  
Db 3 APT 5

RESULT 9  
PT0530  
T-cell receptor beta chain V-D-J region (100-4AK) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0530  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601  
A:Accession: PT0530  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-8 <PEE>  
A:Experimental source: adult thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 32.7%; Score 16; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 APT 4  
|||  
Db 4 APT 6

RESULT 10  
D48186  
ATPase RI subunit - wood tobacco (fragment)  
C:Species: Nicotiana sylvestris (wood tobacco)  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 23-Feb-1997  
C:Accession: D48186  
R:De Paape, R.; Forchioni, A.; Chetrit, P.; Vedel, F.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993  
A:Title: Specific mitochondrial proteins in pollen: presence of an additional ATP synthase  
A:Reference number: A48186; MUID:93317598  
A:Accession: D48186  
A:Status: preliminary  
A:Molecule type: protein  
A:Experimental source: pollen  
A:Note: sequence extracted from NCBI backbone (NCBIP:134871)

Query Match 32.7%; Score 16; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAPTK 5  
|||  
Db 5 LAPYK 9

RESULT 11  
C41978  
calliphramide 3 - bluebottle fly (Calliphora vomitoria)  
C:Species: Calliphora vomitoria  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Mar-1999  
C:Accession: C41978  
R:Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe

Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992  
A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (d)  
A:Reference number: A41978; MUID:92196111  
A:Accession: C41978  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <DUV>  
C:Keywords: amidated carboxyl end; neuropeptide  
F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 32.7%; Score 16; DB 2; Length 9;  
Best Local Similarity 40.0%; Pred. No. 2e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 APTKD 6  
|||  
Db 1 SPSQD 5

RESULT 12  
E41978  
calliphramide 5 - bluebottle fly (Calliphora vomitoria)  
C:Species: Calliphora vomitoria  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Mar-1999  
C:Accession: E41978  
R:Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe  
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992  
A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (d)  
A:Reference number: A41978; MUID:92196111  
A:Accession: E41978  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <DUV>  
C:Keywords: amidated carboxyl end; neuropeptide  
F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 32.7%; Score 16; DB 2; Length 9;  
Best Local Similarity 60.0%; Pred. No. 2e+05;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 APTKD 6  
|||  
Db 1 APQD 5

RESULT 13  
B38740  
Ig kappa chain C region (PY20) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 19-Mar-1997 #sequence\_revision 13-Mar-1998 #text\_change 13-Mar-1998  
C:Accession: B38740  
R:Ruff-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.  
J. Biol. Chem. 266, 6607-6613, 1991  
A:Title: Heavy and light chain variable region sequences and antibody properties of a  
A:Reference number: A38740; MUID:91177923  
A:Accession: B38740  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-9 <RUF>

Query Match 32.7%; Score 16; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 APT 4  
|||  
Db 4 APT 6

RESULT 14

A42266  
peptidylglycine monooxygenase (EC 1.14.17.3), rPAM-5 - rat (fragment)  
N;Alternate names: peptidylglycine alpha-amidating monooxygenase  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-May-1995  
C;Accession: A42266  
R;Eipper, B.A.; Green, C.B.; Campbell, T.A.; Stoffers, D.A.; Keutmann, H.T.; Mains, R.E.  
J. Biol. Chem. 267, 4008-4015, 1992  
A;Title: Alternative splicing and endoproteolytic processing generate tissue-specific fo  
A;Reference number: A42266; MUID:92156145  
A;Accession: A42266  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-9 <EIP>  
A;Experimental source: pituitary  
A;Note: sequence extracted from NCBI backbone (NCBIN:82733, NCBIIP:82750)  
C;Keywords: oxidoreductase

Query Match 32.7%; Score 16; DB 2; Length 9;  
Best Local Similarity 75.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 KDEP 8  
|||  
Db 6 KDTF 9

RESULT 15  
B60110  
repetitive protein antigen 61 - Trypanosoma cruzi (fragment)  
C;Species: Trypanosoma cruzi  
C;Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 18-Jun-1993  
C;Accession: B60110  
R;Hof, D.F.; Kim, K.S.; Otsu, K.; Moser, D.R.; Yost, W.J.; Blumin, J.H.; Donelson, J.E.  
Infect. Immun. 57, 1959-1967, 1989  
A;Title: Trypanosoma cruzi expresses diverse repetitive protein antigens.  
A;Reference number: A60110; MUID:89277508  
A;Accession: B60110  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-6 <HOF>  
A;Note: this is an example of a five residue tandem repeat from this protein; the actual  
C;Keywords: tandem repeat

Query Match 30.6%; Score 15; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 APTK 5  
|||  
Db 2 APKK 5

Search completed: January 17, 2001, 13:43:37  
Job time: 156 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 17, 2001, 13:47:09 ; Search time 20.9 Seconds  
(without alignments)  
13.907 Million cell updates/sec

Title: US-08-765-837-6  
Perfect score: 49  
Sequence: 1 MAPTKDEFD 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 213

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	18	36.7	8	1 UC26_MAIZE	P80632 zea mays (m
2	18	36.7	9	1 DCML_PSECF	P19913 pseudomonas
3	17	34.7	7	1 FAR1_ASCSU	P31889 ascaris suu
4	17	34.7	9	1 YBFR_AZOVI	P25825 azotobacter
5	16	32.7	9	1 FAR3_CALVO	P41858 calliphora
6	16	32.7	9	1 FAR5_CALVO	P41860 calliphora
7	16	32.7	9	1 TKCI_CALVO	P41517 calliphora
8	15	30.6	8	1 CLP_THICU	P80488 thiobacillu
9	15	30.6	9	1 FAR2_CALVO	P41857 calliphora
10	15	30.6	9	1 FAR4_CALVO	P41865 calliphora
11	14	28.6	9	1 FAR4_CALVO	P41859 calliphora
12	13	26.5	9	1 FAR1_CALVO	P41856 calliphora
13	12	24.5	5	1 BIOA_CITFR	P13071 citrobacter
14	12	24.5	5	1 BIOA_SALTY	P12677 salmonella
15	12	24.5	5	1 PRCT_PERAM	P01373 periplaneta
16	12	24.5	6	1 TMOF_SARBU	P41495 sarcophaga
17	12	24.5	7	1 FAR2_ASCSU	P31890 ascaris suu
18	12	24.5	7	1 FARB_CALVO	P41866 calliphora
19	12	24.5	8	1 ACI_THUAL	P18691 thunnus alb
20	12	24.5	8	1 PLP_BRANA	P81707 brassica na
21	12	24.5	9	1 FAR5_ASCSU	P43170 ascaris suu
22	12	24.5	9	1 NSK1_SARBU	P41492 sarcophaga
23	12	24.5	9	1 OXYT_EISRO	P42998 leisenia foe
24	12	24.5	9	1 TRP4_LEUMA	P81736 leucophaea
25	12	24.5	9	1 UF02_MOUSE	P38640 mus musculu
26	12	24.5	9	1 UHA2_HUMAN	P40929 homo sapien
27	12	24.5	9	1 UPA3_HUMAN	P30089 homo sapien
28	11	22.4	3	1 LUXE_VIBFI	P24272 vibrio fisc
29	11	22.4	5	1 BPP7_BOTIN	P30425 bothrops in
30	11	22.4	5	1 SUGA_ACHDO	P19991 acheta dome
31	11	22.4	5	1 TRM3_ECOLI	P13973 escherichia
32	11	22.4	6	1 CIP2_MYTED	P13737 mytilus edu
33	11	22.4	7	1 FAR1_HELTI	P41871 helisoma tr

34 11 22.4 7 1 MSCL\_SALTY p39446 salmonella  
35 11 22.4 7 1 UF03\_MOUSE p38641 mus musculu  
36 11 22.4 8 1 B44K\_PORGI p81886 porphyronon  
37 11 22.4 8 1 RS7\_MYCIT p33564 mycobacteri  
38 11 22.4 9 1 ALI0\_CARMA p81813 carcinus ma  
39 11 22.4 9 1 COXE\_THUOB p80975 thunnus obe  
40 11 22.4 9 1 FARD\_CALVO p41868 calliphora  
41 11 22.4 9 1 HUTU\_KLEAE p12381 klebsiella  
42 11 22.4 9 1 LFCA\_STAAU p36884 staphylococ  
43 11 22.4 9 1 RS11\_SALTY O54296 salmonella  
44 10 20.4 4 1 TUFT\_HUMAN P01858 homo sapien  
45 10 20.4 7 1 FAR4\_PANRE P41875 panagrellus

## ALIGNMENTS

RESULT 1  
UC26\_MAIZE  
ID UC26\_MAIZE STANDARD; PRT; 8 AA.  
AC P80632;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 907)  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=COLEOPTILE;  
RA Touzet P., Ricciardi F., Morin C., Damerval C., Huet J.-C.,  
RA Pernollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
genome analysis program."  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 7.0, ITS MW IS: 57.2 KDA.  
DR MAIZE-ZDPAGE; P80632; COLEOPTILE.  
FT NON\_TER 1  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 990 MW; 9639D6DAB4176B1D CRC64;

Query Match 36.7%; Score 18; DB 1; Length 8;  
Best Local Similarity 42.9%; Pred. No. 8.9e+04;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 2 APTKDEF 8  
Db 1 AEPDQF 7

RESULT 2  
DCML\_PSECF  
ID DCML\_PSECF STANDARD; PRT; 9 AA.  
AC P19913;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-DEC-1992 (Rel. 24, Last annotation update)  
DE CARBON MONOXIDE DEHYDROGENASE LARGE CHAIN (EC 1.2.99.2) (FRAGMENT).  
OS Pseudomonas carboxydoflava.  
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;  
OC Hydrogenophaga.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=90055678; PubMed=2818128;  
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;  
RT "Homology and distribution of CO dehydrogenase structural genes in  
carboxydofrophic bacteria."  
RL Arch. Microbiol. 152:335-341(1989).  
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR -> CO(2) + REDUCED

CC ACCEPTOR.  
CC -1- COFACTOR: MOLYBDENUM.  
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND  
CC SMALL.  
DR PIR: PL0139; PL0139.  
KW Oxidoreductase; Molybdenum.  
FT VARIANT 2 2 N -> M.  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 974 MW; 0224DAB6C2D76DD4 CRC64;

Query Match 36.7%; Score 18; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 APTKD 6  
DB 3 APQVD 7

RESULT 3  
FARL\_ASCSU STANDARD; PRT; 7 AA.

AC P31889;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE FMRFAMIDE-LIKE NEUROPEPTIDE AF1.  
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoidea).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;  
OC Ascarididae; Ascaris.  
RN [1]

RP SEQUENCE.  
RX MEDLINE=90180465; PubMed=2627377;  
RA Cowden C., Stretton A.O.W., Davis R.E.;  
RT "AF1, a sequenced bioactive neuropeptide isolated from the nematode  
RT Ascaris suum";

RL Neuron 2:1465-1473(1989).  
CC -1- FUNCTION: POTENT MODULATOR OF INHIBITORY MOTONEURONS. REDUCES THE  
CC INPUT RESISTANCE AND BLOCKS SLOW OSCILLATORY POTENTIALS IN THESE  
CC CELLS.

CC -1- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF  
CC GANGLIA PARTICULARLY IN THE ANTERIOR REGIONS.  
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.

KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7 AMIDATION  
SQ SEQUENCE 7 AA; 953 MW; 69D40059CBI44350 CRC64;

Query Match 34.7%; Score 17; DB 1; Length 7;  
Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KDEF 8  
DB 1 KNEF 4

RESULT 4  
YBFR\_AZQVI STANDARD; PRT; 9 AA.

AC P25825;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 01-JUL-1993 (Rel. 26, Last annotation update)  
DE HYPOTHETICAL PROTEIN IN BFR 3'REGION (FRAGMENT).  
OS Azotobacter vinelandii.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Azotobacter.  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=92196129; PubMed=1549605;

RA Grossman M.J., Hinton S.M., Minak-Bernero V., Slaughter C.,  
RA Stiefel E.I.;  
RT "Unification of the ferritin family of proteins";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2419-2423(1992).  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M83692; AAA22122.1; -

PIR: B41983; B41983.

KW Hypothetical protein..

FT NON\_TER 9 9

SQ SEQUENCE 9 AA; 947 MW; DF98B5A1B417776D CRC64;

Query Match 34.7%; Score 17; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPTK 5  
DB 1 MAPPR 5

RESULT 5  
FAR3\_CALVO

ID FAR3\_CALVO STANDARD; PRT; 9 AA.

AC P41858;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE CALLIFMRFAMIDE 3.

OS Calliphora vomitoria (Blue blowfly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Oestroidea; Calliphoridae; Calliphora.  
RN [1]

RP SEQUENCE.

RC TISSUE=THORACIC GANGLION;

RX MEDLINE=92196111; PubMed=1549595;

RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,

RA Rehfeld J.F., Thorpe A.;

RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
RT neuropeptides (designated callifmrfa) from the blowfly  
RT Calliphora vomitoria";

CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED  
CC SALIVARY GLAND OF CALLIPHORA.

CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.

PIR: C41978; C41978.

KW Neuropeptide; Amidation.

FT MOD\_RES 9 9 AMIDATION

SQ SEQUENCE 9 AA; 1114 MW; 2F0B0699CAB6C5A7 CRC64;

Query Match 32.7%; Score 16; DB 1; Length 9;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 APTKD 6  
DB 1 SPSQD 5

RESULT 6  
FAR5\_CALVO

ID FAR5\_CALVO STANDARD; PRT; 9 AA.



AC P41860;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CALLIFMRFAMIDE 5.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=THORACIC GANGLION;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
RA Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
neuropeptides (designated callifmrfamides) from the blowfly  
Calliphora vomitoria.";  
RT Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
FAMILY.  
DR PIR: E41978; E41978.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 9  
SQ SEQUENCE 9 AA; 1068 MW; 39D10699CAB6D867 CRC64;  
  
Query Match 32.7%; Score 16; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 APTKD 6  
DB 1 APCQD 5  
  
RESULT 7  
TKCL\_CALVO  
ID TKCL\_CALVO STANDARD; PRT; 9 AA.  
AC P41517;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CALLITACHYKININ I.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RX MEDLINE=95075727; PubMed=7984492;  
RA Lundquist C.T., Clottens F.L., Holman G.M., Nichols R., Nachman R.J.,  
RA Naessel D.R.;  
RT "Callitachykinin I and II, two novel myotropic peptides isolated from  
the blowfly, Calliphora vomitoria, that have resemblances to  
tachykinins.";  
RL Peptides 15:761-768(1994).  
CC -!- FUNCTION: MYOACTIVE PEPTIDE.  
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 9  
SQ SEQUENCE 9 AA; 981 MW; 2417C86B59CDC1B7 CRC64;

Query Match 32.7%; Score 16; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 APT 4  
DB 1 APT 3

RESULT 8  
CLP\_THICU  
ID CLP\_THICU STANDARD; PRT; 8 AA.  
AC P80488;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE CHEMOLITHOTROPH-SPECIFIC PROTEIN (FRAGMENT).  
OS Thiobacillus cuprinus.  
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Thiomonas.  
RN [1]  
RP SEQUENCE.  
RC STRAIN=DSM 5494;  
RA Marin I., Amaro A.M., Jerez C.A., Amils R., Abad J.P.;  
RL Submitted (SEP-1995) to the SWISS-PROT data bank.  
CC -!- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED  
ON CHEMOLITHOTROPHICALLY.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 785 MW; 91487B06DDC2D76D CRC64;

Query Match 30.6%; Score 15; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 APTKDE 7  
DB 1 APVAQE 6

RESULT 9  
FAR2\_CALVO  
ID FAR2\_CALVO STANDARD; PRT; 9 AA.  
AC P41857;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CALLIFMRFAMIDE 2.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=THORACIC GANGLION;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
RA Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
neuropeptides (designated callifmrfamides) from the blowfly  
Calliphora vomitoria.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
CC -!- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED  
SALIVARY GLAND OF CALLIPHORA.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
FAMILY.  
DR PIR: B41978; B41978.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 9  
SQ SEQUENCE 9 AA; 1128 MW; 29D00699CAB6C5A7 CRC64;

Query Match 30.6%; Score 15; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTKD 6  
DB 2 PSQD 5

RESULT 10  
FARA\_CALVO

ID FARA\_CALVO STANDARD; PRT; 9 AA.  
 AC P41855;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE CALLIFMRFAMIDE 10.  
 OS Calliphora vomitoria (Blue blowfly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Calliphoridae; Calliphora.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=THORACIC GANGLION;  
 RX MEDLINE=92196111; PubMed=1549595;  
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
 RA Rehfeld J.F., Thorpe A.;  
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
 RT neuropeptides (designated callifmrFamides) from the blowfly  
 RT Calliphora vomitoria";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR PIR; A44787; A44787.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 1 1 1 AMIDATION.  
 FT UNSURE 1 1 1 OR S OR A.  
 SQ SEQUENCE 9 AA; 1183 MW; 29D00699CAB40457 CRC64;  
  
 Query Match 30.6%; Score 15; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 PTKD 6  
 Db 2 PNRD 5  
  
 RESULT 11  
 FARA\_CALVO STANDARD; PRT; 9 AA.  
 AC P41859;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE CALLIFMRFAMIDE 4.  
 OS Calliphora vomitoria (Blue blowfly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Calliphoridae; Calliphora.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=THORACIC GANGLION;  
 RX MEDLINE=92196111; PubMed=1549595;  
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
 RA Rehfeld J.F., Thorpe A.;  
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
 RT neuropeptides (designated callifmrFamides) from the blowfly  
 RT Calliphora vomitoria";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR PIR; D41978; D41978.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 1182 MW; 31730699CAB6D457 CRC64;  
  
 Query Match 28.6%; Score 14; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 PTKD 6

Db 2 PNQD 5  
  
 RESULT 12  
 FARA\_CALVO STANDARD; PRT; 9 AA.  
 ID FARA\_CALVO STANDARD; PRT; 9 AA.  
 AC P41856;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE CALLIFMRFAMIDE 1.  
 OS Calliphora vomitoria (Blue blowfly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Calliphoridae; Calliphora.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=THORACIC GANGLION;  
 RX MEDLINE=92196111; PubMed=1549595;  
 RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
 RA Rehfeld J.F., Thorpe A.;  
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
 RT neuropeptides (designated callifmrFamides) from the blowfly  
 RT Calliphora vomitoria";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
 CC -!- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED  
 CC SALIVARY GLAND OF CALLIPHORA.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR PIR; A41978; A41978.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 1169 MW; 29D00699CAB6C6C7 CRC64;  
  
 Query Match 26.5%; Score 13; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 PTKD 6  
 Db 2 PQOD 5  
  
 RESULT 13  
 BIOA\_CITFR STANDARD; PRT; 5 AA.  
 ID BIOA\_CITFR STANDARD; PRT; 5 AA.  
 AC P13071;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE  
 DE (EC 2.6.1.62) (7,8-DIAMINO-7-AMINO-7-OXONONANOATE AMINOTRANSFERASE) (FRAGMENT).  
 DE AMINOTRANSFERASE)  
 GN BIOA.  
 OS Citrobacter freundii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Citrobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89006280; PubMed=2971595;  
 RA Shivan D., Campbell A.;  
 RT "Transcriptional regulation and gene arrangement of Escherichia coli,  
 RT Citrobacter freundii and Salmonella typhimurium biotin operons";  
 RL Gene 67:203-211(1988).  
 CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 8-AMINO-7-  
 CC OXONONANOATE = S-ADENOSYL-4-METHYLTHIO-2-OXOBUTANOATE +  
 CC 7,8-DIAMINONONANOATE.  
 CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -!- PATHWAY: BIOTIN BIOSYNTHESIS.  
 CC -!- SUBUNIT: HOMODIMER.  
 CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT

```

CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21922; CAB25179.1; -.
DR INTERPRO; IPR000954; -.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
KW Biotin biosynthesis; Transferase; Amino transferase;
KW Pyridoxal phosphate. 5
FT NON_TER 5
SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1A6F00000 CRC64;

Query Match 24.5%; Score 12; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TKDE 7
Db 2 TTDD 5

RESULT 14
BIOA_SALTY STANDARD; PRT; 5 AA.
ID BIOA_SALTY STANDARD; PRT; 5 AA.
AC P12677;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE
DE (EC 2.6.1.62) (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA
DE AMINOTRANSFERASE) (FRAGMENT).
GN BIOA.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89006280; PubMed=2971595;
RA Shuan D., Campbell A.;
RT "Transcriptional regulation and gene arrangement of Escherichia coli,
RT Citrobacter freundii and Salmonella typhimurium biotin operons.";
RL Gene 67:203-211(1988).
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + 8-AMINO-7-
CC OXONONANOATE = S-ADENOSYL-4-METHYLTHIO-2-OXOBUTANOATE +
CC 7,8-DIAMINONONANOATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: BIOTIN BIOSYNTHESIS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21923; -. NOT_ANNOTATED_CDS.
DR STYGENE; SGI0026; BIOA.
DR INTERPRO; IPR000954; -.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
KW Biotin biosynthesis; Transferase; Amino transferase;
KW Pyridoxal phosphate. 5
FT NON_TER 5
SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1A6F00000 CRC64;

```

```

SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1A6F00000 CRC64;

Query Match 24.5%; Score 12; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TKDE 7
Db 2 TTDD 5

RESULT 15
PRTCT_PERAM STANDARD; PRT; 5 AA.
ID PRTCT_PERAM STANDARD; PRT; 5 AA.
AC P01373;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE PROCTOLIN.
OS Periplaneta americana (American cockroach),
OS Limulus polyphemus (Atlantic horseshoe crab), and
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattodea; Blattidae; Periplaneta.
RN [1]
RP SEQUENCE.
RC SPECIES=P-AMERICANA;
RX MEDLINE=76074708; PubMed=576;
RA Starratt A.N., Brown B.E.;
RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter
RT in insects.";
RL Life Sci. 17:1253-1256(1975).
RN [2]
RP BIOLOGICAL SOURCE.
RC SPECIES=P-AMERICANA;
RX MEDLINE=81225865; PubMed=6113690;
RA O'Shea M., Adams M.E.;
RT "Pentapeptide (proctolin) associated with an identified neuron.";
RL Science 213:567-569(1981).
RN [3]
RP SEQUENCE.
RC SPECIES=L-POLYPHEMUS;
RX MEDLINE=90287800; PubMed=2356151;
RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA Shabanowitz J.;
RT "Identification of proctolin in the central nervous system of the
RT horseshoe crab, Limulus polyphemus.";
RL Peptides 11:205-211(1990).
RN [4]
RP SEQUENCE.
RC SPECIES=C-MAENAS;
RX MEDLINE=86232789; PubMed=2872661;
RA Stangier J., Dirksen H., Keller R.;
RT "Identification and immunocytochemical localization of proctolin in
RT pericardial organs of the shore crab, Carcinus maenas.";
RL Peptides 7:67-72(1986).
CC -!- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY.
CC MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
CC -!- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN
CC THE CRAB PERICARDIAL ORGANS.
DR PIR; A01644; HOROHA.
DR PIR; A60411; A60411.
KW Neuropeptide.
SQ SEQUENCE 5 AA; 649 MW; 71B7673B44600000 CRC64;

Query Match 24.5%; Score 12; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 3 PT 4  
| |  
Db 4 PT 5

Search completed: January 17, 2001, 13:47:10  
Job time: 279 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:44:40 ; Search time 58.21 Seconds  
(without alignments)  
18.122 Million cell updates/sec

Title: US-08-765-837-6  
Perfect score: 49  
Sequence: 1 MAPTKDEFD 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 467

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	18	36.7	9	12	Q70140 human immun
2	17	34.7	8	8	Q9T2Y3 begonia for
3	16	32.7	8	5	Q94623 manduca sex
4	16	32.7	9	2	Q9P735 streptomyce
5	15	30.6	8	6	O18854 canis fami
6	15	30.6	8	10	Q42507 triticum ae
7	14	28.6	8	4	Q9P0K3 homo sapien
8	14	28.6	8	5	Q9N6M5 toxoplasma
9	14	28.6	8	12	Q9J205 hepatitis c
10	14	28.6	9	2	P72345 pseudomonas
11	13	26.5	8	2	Q9R3X0 planktothri
12	13	26.5	8	2	Q9R5L7 clostridium
13	13	26.5	8	4	Q15889 homo sapien
14	13	26.5	8	8	Q9TMN4 begonia for
15	13	26.5	8	8	Q9P2Y2 begonia tai
16	13	26.5	8	8	Q9P2Y1
17	13	26.5	8	10	Q9XGL9 begonia nan
18	13	26.5	8	10	Q9XGL8 begonia chi
19	13	26.5	8	10	Q9STD5 begonia pal

20	13	26.5	9	3	Q9UR18
21	13	26.5	9	3	Q9P8E5
22	13	26.5	9	7	Q78225
23	13	26.5	9	7	Q78226
24	13	26.5	9	8	Q9TNF6
25	13	26.5	9	8	Q9TNF5
26	13	26.5	9	8	Q9TNF4
27	13	26.5	9	8	Q9TNF3
28	13	26.5	9	8	Q9TNF2
29	13	26.5	9	8	Q9TNF1
30	13	26.5	9	8	Q9TNF0
31	13	26.5	9	8	Q9TNE9
32	13	26.5	9	8	Q9TNE8
33	13	26.5	9	8	Q9TNE7
34	13	26.5	9	8	Q9TNE6
35	13	26.5	9	8	Q9TNE5
36	13	26.5	9	8	Q9TNE4
37	13	26.5	9	8	Q9TNE3
38	13	26.5	9	8	Q9TNE2
39	13	26.5	9	8	Q9TNE1
40	13	26.5	9	8	Q9TNE0
41	13	26.5	9	8	Q9TND9
42	13	26.5	9	8	Q9TND8
43	13	26.5	9	8	Q9TND7
44	13	26.5	9	8	Q9TKF2
45	13	26.5	9	8	Q9THM0

#### ALIGNMENTS

RESULT 1

Q70140

ID Q70140 PRELIMINARY: PRT: 9 AA.

AC Q70140;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE TAT PROTEIN (FRAGMENT).

GN TAT.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentiviruses.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=020;

RX MEDLINE=95194694; PubMed=7888189;

RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L., Hahn B.H.;

RA McCutchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.;

RT "Genetic variation of HIV type 1 in four World Health Organization-

RT sponsored vaccine evaluation sites; generation of functional envelope-

RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,

RT and E. WHO Network for HIV Isolation and Characterization.";

RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=020;

RX MEDLINE=96190564; PubMed=8627686;

RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,

RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B., Briesen H.,

RA Beddows S., Weber J., Sharp P.M., Shaw G.M., Hahn B.H.;

RT "Molecular cloning and analysis of functional envelope genes from

RT human immunodeficiency virus type 1 sequence subtypes A through G. The

RT WHO and NIAID Networks for HIV Isolation and Characterization.";

RL J. Virol. 70:1651-1657(1996).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=020;

RA Allen E.E.;

RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL; U08794; AAB05175.1; -.

FT NON\_TER 1 1

SQ SEQUENCE 9 AA; 1098 MW; 5B76D40ABIAB01A3 CRC64;

Query Match 36.7% Score 18; DB 12; Length 9;  
Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 DEFD 9  
| | |  
Db 6 DRED 9

RESULT 2  
Q9T2Y3 PRELIMINARY; PRT; 8 AA.  
AC Q9T2Y3;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE RBCL PROTEIN (FRAGMENT).  
GN RBCL.  
OS Begonia formosana.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
OC Cucurbitales; Begoniaceae; Begonia.  
OX NCBI\_TaxID=80370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chiang T.Y.;  
RT "Sequence announcement: rbcL promotor gene in Begonia formosana.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ009599; CAB52123.1; -  
DR EMBL: AJ009598; CAB52121.1; -  
KW Chloroplast.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 907 MW; FA21AAB1B6C775B6 CRC64;

Query Match 34.7% Score 17; DB 8; Length 8;  
Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
Matches 1; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAPTKD 6  
| | |  
Db 1 MSPQTD 6

RESULT 3  
Q94623 PRELIMINARY; PRT; 8 AA.  
AC Q94623;  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE MSUSP-2 PROTEIN (FRAGMENT).  
GN USP.  
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=EPIDERMIS, WING;  
RX MEDLINE=97165493; PubMed=9013254;  
RA Jindra M., Huang J.Y., Malone F., Asahina M., Riddiford L.M.;  
RT "Identification and mRNA developmental profiles of two ultraspiracle isoforms in the epidermis and wings of Manduca sexta.";  
RL Insect Mol. Biol. 6:41-53(1997).  
DR EMBL: U57921; AAB64235.1; -  
FT NON\_TER 6  
SQ SEQUENCE 8 AA; 892 MW; F165BB0415A76B16 CRC64;

Query Match 32.7% Score 16; DB 5; Length 8;  
Best Local Similarity 33.3%; Pred. No. 3.7e+05;  
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAPTKD 6  
| | |  
Db 1 MEPSRE 6

RESULT 4  
Q9R735 PRELIMINARY; PRT; 9 AA.  
AC Q9R735;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE FKBA PROTEIN (FRAGMENT).  
GN FKBA.  
OS Streptomyces chrysomallus.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1899;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94341259; PubMed=8062824;  
RA Pahl A., Keller U.;  
RT "Streptomyces chrysomallus FKBP-33 is a novel immunophilin consisting of two FK506 binding domains; its gene is transcriptionally coupled to the FKBP-12 gene.";  
RL EMBL: J. 13:3472-3480(1994).  
DR EMBL: Z34523; CAA84282.1; -  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1047 MW; 9A8BCB07633B1045 CRC64;

Query Match 32.7% Score 16; DB 2; Length 9;  
Best Local Similarity 44.4%; Pred. No. 3.7e+05;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MAPTKDEFD 9  
| | | |  
Db 1 MSIEKPEVD 9

RESULT 5  
O18854 PRELIMINARY; PRT; 8 AA.  
AC O18854;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE CD19 ANTIGEN (FRAGMENT).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Liu P.-C., Shibuya H., Katz M.L., Johnson G.S.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF024717; AAB81967.1; -  
FT NON\_TER 1  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 832 MW; 6735A1ADDB1325A7 CRC64;

Query Match 30.6% Score 15; DB 6; Length 8;  
Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PTKD 6  
| | |

```
Db 1 PSKE 4

RESULT 6
Q42507 PRELIMINARY; PRT; 8 AA.
AC Q42507;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HEAT SHOCK PROTEIN (FRAGMENT).
GN HSP70C OR HSP70A OR HSP70B.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV MUSTANG;
RX MEDLINE=96189275; PubMed=8605312;
RA Joshi C.P., Kumar S., Nguyen H.T.;
RT "Application of modified differential display technique for cloning
and sequencing of the 3' region from three putative members of wheat
HSP70 gene family.";
RL Plant Mol. Biol. 30:641-646(1996).
DR EMBL; L41507; AAB02333.1; -
DR EMBL; L41505; AAB02331.1; -
DR EMBL; L41506; AAB02332.1; -
KW Heat shock..
FT NON_TER 1
SQ SEQUENCE 8 AA; 886 MW; 71B2CB1B10532768 CRC64;

Query Match 30.6%; Score 15; DB 10; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.7e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PTKDEFD 9
| : | |
DB 2 PKIEVD 8

RESULT 7
Q9P0K3 PRELIMINARY; PRT; 8 AA.
AC Q9P0K3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CGMP-SPECIFIC PHOSPHODIESTERASE PDE5A2 (FRAGMENT).
GN PDE5A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20145478; PubMed=10679249;
RA Lin C.S., Lau A., Tu R., Lue T.F.;
RT "Identification of three alternative first exons and an intronic
promoter of human PDE5A gene.";
RL Biochem. Biophys. Res. Commun. 268:596-602(2000).
DR EMBL; AF155195; AAF40302.1; -
FT NON_TER 8
SQ SEQUENCE 8 AA; 908 MW; E8A33AA879D76726 CRC64;

Query Match 28.6%; Score 14; DB 4; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.7e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPTKDE 7
| | | |

Db 1 MLPEGDK 7

RESULT 8
Q9NGM5 PRELIMINARY; PRT; 8 AA.
AC Q9NGM5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE DIHYDROFOLATE REDUCTASE THYMIDYLATE SYNTHASE (FRAGMENT).
GN FOL1.
OS Toxoplasma gondii.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida; Sarcocystidae;
OC Toxoplasma.
OX NCBI_TaxID=5811;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RH, COUGAR TC751G34, SEA OTTER TC828G1, AND BEVERLEY;
RA Lehmann T., Blackston C.R., Parmley S.F., Remington J.S., Dubey J.P.;
RT "Strain Typing of Toxoplasma gondii: Comparison of Antigen-Coding and
Housekeeping Genes.";
RL J. Parasitol. 0:0-0(2000).
DR EMBL; AF249695; AAF79153.1; -
DR EMBL; AF249692; AAF79150.1; -
DR EMBL; AF249693; AAF79151.1; -
DR EMBL; AF249694; AAF79152.1; -
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 1035 MW; 33CAAAA05B133044 CRC64;

Query Match 28.6%; Score 14; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 DEF 8
| : |
DB 6 DDF 8

RESULT 9
Q9J205 PRELIMINARY; PRT; 8 AA.
AC Q9J205;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE TRUNCATED POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HC41;
RA Fan X., DiBisceglie A.M.;
RT "Identification of liver-specific quasisppecies of the hepatitis C
virus in chronically infected patients.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211054; AAF30114.1; -
KW Polyprotein.
FT NON_TER 1
SQ SEQUENCE 8 AA; 917 MW; B1D41AF7776DCA CRC64;

Query Match 28.6%; Score 14; DB 12; Length 8;
Best Local Similarity 42.9%; Pred. No. 3.7e+05;
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 APTKDEF 8
| | |
DB 2 APPXXRF 8
```

```

RESULT 10
P72345
ID P72345 PRELIMINARY; PRT; 9 AA.
AC P72345;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 1.0 KDA PROTEIN (FRAGMENT).
GN TABA.
OS Pseudomonas syringae.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93167809; PubMed=7679566;
RA Barta T.M., Kinscherf T.G., Uchytel T.F., Willis D.K.;
RT "DNA sequence and transcriptional analysis of the tbla gene required
RT for tabtoxin biosynthesis by Pseudomonas syringae.";
RL Appl. Environ. Microbiol. 59:458-466(1993).
DR EMBL; S54909; AAB5381.2; -.
KW Hypothetical protein.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1037 MW; 2B34D9D5BB05B047 CRC64;

Query Match 28.6%; Score 14; DB 2; Length 9;
Best Local Similarity 33.3%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PTKDEF 8
Db 2 PISESF 7

RESULT 11
Q9R3X0
ID Q9R3X0 PRELIMINARY; PRT; 8 AA.
AC Q9R3X0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT
DE (FRAGMENT).
GN RECL.
OS Planktothrix rubescens.
OC Bacteria; Cyanobacteria; Oscillatoriales; Planktothrix.
OX NCBI_TaxID=59512;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BC-PLA 9316, AND BC-PLA 9303;
RA Beard S.J., Handley B.A., Hayes P.K., Walsby A.E.;
RT "The diversity of gas vesicle genes in Planktothrix rubescens from
RT Lake Zurich.";
RL Microbiology 145:2757-2768(1999).
DR EMBL; A1132249; CAB59537.1; -.
DR EMBL; A1132248; CAB59534.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 957 MW; 33D1AAA685BB19CB CRC64;

```

```

Query Match 26.5%; Score 13; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 EFD 9
Db 1 EFE 3

RESULT 12

```

```

Q9R5L7
ID Q9R5L7 PRELIMINARY; PRT; 8 AA.
AC Q9R5L7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE 1,4-BETA-D-GLUCAN GLUCANOHYDROLASE (EC 3.2.1.4) (FRAGMENT).
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE.
RX MEDLINE=92231850; PubMed=1567379;
RA Romaniec M.P., Fauth U., Kobayashi T., Huskisson N.S., Barker P.J.,
RA Demain A.L.;
RT "Purification and characterization of a new endoglucanase from
RT Clostridium thermocellum.";
RL Biochem. J. 283:69-73(1992).
SQ SEQUENCE 8 AA; 823 MW; C2CIAB1DD9D1B775 CRC64;

Query Match 26.5%; Score 13; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 APT 4
Db 1 SPT 3

RESULT 13
Q15889
ID Q15889 PRELIMINARY; PRT; 8 AA.
AC Q15889;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE (CLONE XP15H8B) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coombebaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.Y.,
RA Caskey C.T.H.;
RL Hum. Mol. Genet. 0:0-0(0).
DR EMBL; L32070; AAA73879.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 865 MW; 0474472325A761E7 CRC64;

Query Match 26.5%; Score 13; DB 4; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.7e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTK 5
Db 3 PSK 5

RESULT 14
Q9TMN4
ID Q9TMN4 PRELIMINARY; PRT; 8 AA.
AC Q9TMN4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RBCL PROTEIN (FRAGMENT).

```



GN RBCL.  
 OS Begonia formosana.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 CC Cucurbitales; Begoniaceae; Begonia.  
 OX NCBI\_TaxID=80370;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chiang T.Y.;  
 RT "Sequence announcement: rbcL promotor gene in Begonia formosana.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ009597; CAB52119.1; -  
 KW Chloroplast.  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 921 MW; FA21AB01B6C775B6 CRC64;

Query Match 26.5%; Score 13; DB 8; Length 8;  
 Best Local Similarity 66.7%; Pred. NO. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAP 3  
 I:I  
 Db 1 MSP 3

## RESULT 15

Q9T2Y2 PRELIMINARY; PRT; 8 AA.  
 AC Q9T2Y2;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE RBCL PROTEIN (FRAGMENT).  
 GN RBCL.  
 OS Begonia taipeiensis.  
 OG Chloroplast.  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 CC Cucurbitales; Begoniaceae; Begonia.  
 OX NCBI\_TaxID=80371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chiang T.Y.;  
 RT "Sequence announcement: rbcL promotor gene in Begonia formosana.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ009602; CAB52129.1; -  
 DR EMBL; AJ009600; CAB52125.1; -  
 DR EMBL; AJ009601; CAB52127.1; -  
 KW Chloroplast.  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 921 MW; FA21AB01B6C775B6 CRC64;

Query Match 26.5%; Score 13; DB 8; Length 8;  
 Best Local Similarity 66.7%; Pred. NO. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAP 3  
 I:I  
 Db 1 MSP 3

Search completed: January 17, 2001, 13:44:41  
 Job time: 195 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:42:56 ; Search time 31.16 Seconds  
(without alignments)  
5.187 Million cell updates/sec

Title: US-08-765-837-6  
Perfect score: 49  
Sequence: 1 MAPTKDEFD 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues  
Total number of hits satisfying chosen parameters: 44977

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/1aa/6\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/1aa/PCITUS\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	25	51.0	7	2	US-08-637-759B-462
2	25	51.0	7	3	US-08-871-355A-462
3	23	46.9	7	5	5194592-48
4	23	46.9	8	2	US-08-318-856A-4
5	23	46.9	8	3	US-09-075-257A-9
6	22	44.9	8	1	US-08-068-395A-5
7	22	44.9	8	1	US-08-464-365-5
8	22	44.9	8	1	US-08-462-894-23
9	22	44.9	8	1	US-08-206-185-23
10	22	44.9	8	2	US-08-318-856A-3
11	22	44.9	8	3	US-08-335-733D-39
12	22	44.9	8	3	US-08-335-733D-40
13	22	44.9	8	3	US-08-335-733D-41
14	22	44.9	8	3	US-08-335-733D-42
15	22	44.9	8	3	US-08-582-776C-27
16	22	44.9	8	3	US-08-434-831B-27
17	22	44.9	8	4	PCT-US95-02121-17
18	22	44.9	9	2	US-08-706-741B-49
19	22	44.9	9	2	US-08-318-856A-20
20	22	44.9	9	2	US-08-318-856A-27
21	22	44.9	9	2	US-08-924-695A-49
22	21	42.9	4	5	5210029-5
23	21	42.9	4	5	5256769-4
24	21	42.9	4	5	5464939-5
25	21	42.9	5	1	US-08-335-198-41
26	21	42.9	6	1	US-08-240-514-7
27	21	42.9	6	2	US-08-612-302A-7
28	21	42.9	7	1	US-07-958-903A-24

29	21	42.9	7	1	US-07-800-364B-3
30	21	42.9	7	1	US-08-462-018-24
31	21	42.9	7	1	US-08-823-245-24
32	21	42.9	7	2	US-08-687-958A-19
33	21	42.9	7	4	PCT-US91-03388-3
34	21	42.9	7	4	PCT-US92-09443A-24
35	21	42.9	7	5	5179007-7
36	21	42.9	7	5	5194592-68
37	21	42.9	8	1	US-07-958-903A-54
38	21	42.9	8	1	US-08-462-018-54
39	21	42.9	8	1	US-08-189-331-170
40	21	42.9	8	1	US-08-823-245-54
41	21	42.9	8	2	US-08-471-068-170
42	21	42.9	8	3	US-08-335-733D-43
43	21	42.9	8	4	PCT-US92-09443A-54
44	21	42.9	9	1	US-07-958-903A-1
45	21	42.9	9	1	US-07-958-903A-25

ALIGNMENTS

RESULT 1  
US-08-637-759B-462  
; Sequence 462, Application US/08637759B  
; Patent No. 5876931  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; STREET: 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/637,759B  
; FILING DATE: 03-MAY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/02875  
; FILING DATE: 11-DEC-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPMS 101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 462:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; US-08-637-759B-462

Query Match 51.0%; Score 25; DB 2; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0;

QY 5 KDEFD 9  
|:|:|  
Db 1 KDEFD 5

## RESULT 2

US-08-871-355A-462  
; Sequence 462, Application US/08871355A  
; Patent No. 6015669  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; STREET: 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871.355A  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/02875  
; FILING DATE: 11-DEC-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPMS 101 CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 462:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-08-871-355A-462

Query Match 51.0%; Score 25; DB 3; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KDEFD 9  
|:|:|  
Db 1 KDEFD 5

## RESULT 3

US-08-871-355A-462  
; Patent No. 5194592  
; APPLICANT: YOSHIDA, HAJIME  
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO NOVEL  
; POLYPEPTIDES DERIVATIVES OF HUMAN GRANULOCYTE COLONY  
; STIMULATING FACTOR  
; NUMBER OF SEQUENCES: 83  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/318.527  
; FILING DATE: 3-MAR-1989  
; PRIOR APPLICATION DATA:

Query Match 46.9%; Score 23; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

; APPLICATION NUMBER: 136,647  
; FILING DATE: 22-DEC-1987  
; SEQ ID NO:48:  
; LENGTH: 7  
5194592-48

Query Match 46.9%; Score 23; DB 5; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPTK 5  
|:|:|  
Db 1 MAPTR 5

## RESULT 4

US-08-318-856A-4  
; Sequence 4, Application US/08318856A  
; Patent No. 5972351  
; GENERAL INFORMATION:  
; APPLICANT: Adrian V.S. Hill, et al.  
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-  
; TITLE OF INVENTION: RESTRICTED CTL EPTIOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1+  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,856A  
; FILING DATE: October 3, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 92 08 068.8  
; FILING DATE: April 3, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 92 17 704.7  
; FILING DATE: August 20, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/GB93/00711  
; FILING DATE: April 5, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee Cheng  
; REGISTRATION NUMBER: 40,949  
; REFERENCE/DOCKET NUMBER: 263-PP1R157705  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEFAX: (202) 721-8250  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acid residues  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium falciparum  
US-08-318-856A-4

Query Match 46.9%; Score 23; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TKDEFD 9  
:||||  
Db 2 SKDELD 7

RESULT 5  
US-09-075-257A-9  
; Sequence 9, Application US/09075257A  
; Patent No. 6074645  
; GENERAL INFORMATION:  
; APPLICANT: DIAMOND, DON JEFFREY  
; APPLICANT: YOKO, JOANNE  
; TITLE OF INVENTION: IMMUNO-REACTIVE PEPTIDE CTL EPITOPES  
; TITLE OF INVENTION: OF HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BART G. NEWLAND  
; STREET: 555 13TH STREET, NW SUITE 701E  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,257A  
; FILING DATE: 11-MAY-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/021,298  
; FILING DATE: 10-FEB-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/950,064  
; FILING DATE: 14-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/747,488  
; FILING DATE: 12-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NEWLAND, BART G  
; REGISTRATION NUMBER: 31,282  
; REFERENCE/DOCKET NUMBER: 1954-112CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-09-075-257A-9

Query Match 46.9%; Score 23; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 PTKD 6  
:||||  
Db 2 PTKD 5

RESULT 6  
US-08-068-395A-5  
; Sequence 5, Application US/08068395A  
; Patent No. 5496719  
; GENERAL INFORMATION:  
; APPLICANT: YAMADA, YUKIO

; APPLICANT: ASAMI, OSAMU  
; APPLICANT: SUGIYAMA, HIDEHIKO  
; APPLICANT: IDEKABA, CHIE  
; APPLICANT: HOSHINO, FUMIHIKO  
; APPLICANT: HIRAI, MASANA  
; APPLICANT: KAJINO, TSUTOMU  
; APPLICANT: IMAEDA, TAKAO  
; APPLICANT: SARAI, KIYOKO  
; TITLE OF INVENTION: POLYPEPTIDE POSSESSING PROTEIN DISULFIDE  
; TITLE OF INVENTION: ISOMERASE ACTIVITY GENE ENCODING THE SAME AND PROCESS FOR  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/068,395A  
; FILING DATE: 19930527  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 4-135254  
; FILING DATE: 27-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-44013  
; FILING DATE: 04-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-44014  
; FILING DATE: 04-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5496719man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 68-228-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-068-395A-5

Query Match 44.9%; Score 22; DB 1; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0;  
Qy 5 KDEFD 9  
:||||  
Db 1 KDTEFD 5

RESULT 7  
US-08-464-365-5  
; Sequence 5, Application US/08464365  
; Patent No. 5700659  
; GENERAL INFORMATION:  
; APPLICANT: YAMADA, YUKIO  
; APPLICANT: ASAMI, OSAMU  
; APPLICANT: SUGIYAMA, HIDEHIKO  
; APPLICANT: IDEKABA, CHIE

APPLICANT: HOSHINO, FUMIHIKO  
APPLICANT: HIRAI, MASANA  
APPLICANT: KAJINO, TSUTOMU  
APPLICANT: IMAEDA, TAKAO  
APPLICANT: SARAI, KIYOKO  
TITLE OF INVENTION: POLYPEPTIDE POSSESSING PROTEIN DISULFIDE  
TITLE OF INVENTION: ISOMERASE ACTIVITY GENE ENCODING THE SAME AND PROCESS FOR  
TITLE OF INVENTION: PRODUCING THE SAME  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLOM, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,365  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-135254  
FILING DATE: 27-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-44013  
FILING DATE: 04-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-44014  
FILING DATE: 04-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: ObloM, No. 5700659man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 68-228-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-365-5

Query Match 44.9%; Score 22; DB 1; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KDEFD 9  
|||  
Db 1 KDTFD 5

## RESULT 8

US-08-462-894-23  
; Sequence 23, Application US/08462894  
; Patent No. 572312  
; GENERAL INFORMATION:  
; APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE  
; APPLICANT: HAENDLER, BERNHARD  
; APPLICANT: KRAETZSCHMAR, JOERN  
; APPLICANT: SCHLEUNING, WOLF-DIETER  
; APPLICANT: ALAGON, ALEJANDRO  
; APPLICANT: POSSANI, LOURIVAL  
; APPLICANT: CUEVAS-AGUIRRE, DELIA

TITLE OF INVENTION: COLLAGEN-INDUCED PLATELET AGGREGATION  
TITLE OF INVENTION: INHIBITOR  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,894  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/206,185  
FILING DATE: 07-MAR-1994  
APPLICATION NUMBER: US 08/116,889  
FILING DATE: 07-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/EP92/02052  
FILING DATE: 04-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/914,383  
FILING DATE: 17-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/814,884  
FILING DATE: 31-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/756,211  
FILING DATE: 05-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: SCH 1359  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-462-894-23

Query Match 44.9%; Score 22; DB 1; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 PTKDEFD 9  
|||  
Db 2 PPGDNFD 8

## RESULT 9

US-08-206-185-23  
; Sequence 23, Application US/08206185  
; Patent No. 5756454  
; GENERAL INFORMATION:  
; APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE  
; APPLICANT: HAENDLER, BERNHARD  
; APPLICANT: KRAETZSCHMAR, JOERN  
; APPLICANT: SCHLEUNING, WOLF-DIETER  
; APPLICANT: ALAGON, ALEJANDRO  
; APPLICANT: POSSANI, LOURIVAL

APPLICANT: CUEVAS-AGUIRRE, DELIA  
 TITLE OF INVENTION: COLLAGEN-INDUCED PLATELET AGGREGATION  
 NUMBER OF SEQUENCES: 36  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MILLER, WHITE, ZELANO & BRANIGAN, P.C.  
 STREET: 2200 CLARENDON BLVD., SUITE 1400  
 CITY: ARLINGTON  
 STATE: VIRGINIA  
 COUNTRY: US  
 ZIP: 22201  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/206,185  
 FILING DATE: 07-MAR-1994  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/116,889  
 FILING DATE: 07-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/EP92/02052  
 FILING DATE: 04-SEP-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/914,383  
 FILING DATE: 17-JUL-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/814,884  
 FILING DATE: 31-DEC-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/756,211  
 FILING DATE: 05-SEP-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: HAMLET-KING, DIANA  
 REGISTRATION NUMBER: 33,302  
 REFERENCE/DOCKET NUMBER: SCH 1359  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-243-6333  
 TELEFAX: 703-243-6410  
 INFORMATION FOR SEQ ID NO: 23:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-206-185-23

Query Match 44.9%; Score 22; DB 1; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 1.3e+05;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0;

QY 3 PTKDEF 9  
 Db 2 PPGDNF 8

RESULT 10  
 US-08-318-856A-3  
 Sequence 3, Application US/08318856A  
 Patent No. 5972351  
 GENERAL INFORMATION:  
 APPLICANT: Adrian V.S. Hill, et al.  
 TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-  
 RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE  
 ANTIGENS (AS AMENDED)  
 NUMBER OF SEQUENCES: 86  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

STREET: 2033 K Street, N.W., Suite 800  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Wordperfect 5.1+  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/318,856A  
 FILING DATE: October 3, 1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 92 08 068.8  
 FILING DATE: April 3, 1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 92 17 704.7  
 FILING DATE: August 20, 1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/GB93/00711  
 FILING DATE: April 5, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lee Cheng  
 REGISTRATION NUMBER: 40,949  
 REFERENCE/DOCKET NUMBER: 263-PFIR15770US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 721-8200  
 TELEFAX: (202) 721-8250  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8 amino acid residues  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ORIGINAL SOURCE:  
 ORGANISM: Plasmodium falciparum  
 US-08-318-856A-3

Query Match 44.9%; Score 22; DB 2; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0;

QY 5 KDEFD 9  
 Db 3 KDELD 7

RESULT 11  
 US-08-335-733D-39  
 Sequence 39, Application US/08335733D  
 Patent No. 6042831  
 GENERAL INFORMATION:  
 APPLICANT: Beretta, Alberto  
 TITLE OF INVENTION: HIV PROTEIN EPITOPES  
 IMMUNOLOGICALLY HOMOLOGOUS TO HLA  
 NUMBER OF SEQUENCES: 90  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Baker & Botts, L.L.P.  
 STREET: 30 Rockefeller Plaza  
 CITY: New York  
 STATE: NY  
 COUNTRY: U.S.A.  
 ZIP: 10112-0228  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/335,733D

FILING DATE: 10-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MacLeod, Janet M  
REGISTRATION NUMBER: 35,263  
REFERENCE/DOCKET NUMBER: A39928-PCT-USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-408-2500  
TELEFAX: 212-765-2519  
TELEX:

Query Match 44.9%; Score 22; DB 3; Length 8;  
Best Local Similarity 80.0%; Pred. NO. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels

```

RESULT 12
US-08-335-733D-40
: Sequence 40, Application US/08335733D
: Patent No. 6042831
: GENERAL INFORMATION:
:
: APPLICANT: Beretta, Alberto
: TITLE OF INVENTION: HIV PROTEIN EST
: TITLE OF INVENTION: IMMUNOLOGICAL
: NUMBER OF SEQUENCES: 90
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Baker & Botts, L.L.P
: STREET: 30 Rockefeller Plaza
: CITY: New York
: STATE: NY
: COUNTRY: U.S.A.
: ZIP: 10112-0228
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/335,733
: FILING DATE: 10-NOV-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: MacLeod, Janet M
: REGISTRATION NUMBER: 35,263
: REFERENCE/DOCKET NUMBER: A29928
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-408-2500
: TELEFAX: 212-765-2519
:
: INFORMATION FOR SEQ ID NO: 40:

```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-335-733D-40

```

Query Match 44.9%; Score 22; DB 3; Length 8;  
Best Local Similarity 80.0%; Pred. NO. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels

RESULT , 13  
US-08-335-733D-41  
; Sequence 41, Application US/08335733D  
; Patent No. 6042831

Query Match 44.58; Score 22; DB 3; Length 8;



Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPTK 5  
:||||  
Db 2 VAPTK 6

## RESULT 14

US-08-335-733D-42  
; Sequence 42, Application US/08335733D  
; Patent No. 6042831  
; GENERAL INFORMATION:  
; APPLICANT: Beretta, Alberto  
; TITLE OF INVENTION: HIV PROTEIN EPITOPES  
; TITLE OF INVENTION: IMMUNOLOGICALLY HOMOLOGOUS TO HLA  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Baker & Botts, L.L.P.  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 10112-0228  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/335.733D  
; FILING DATE: 10-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MacLeod, Janet M  
; REGISTRATION NUMBER: 35,263  
; REFERENCE/DOCKET NUMBER: A29928-PCT-USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-408-2500  
; TELEFAX: 212-765-2519  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-08-335-733D-42

Query Match 44.9%; Score 22; DB 3; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPTK 5  
:||||  
Db 1 VAPTK 5

## RESULT 15

US-08-582-776C-27  
; Sequence 27, Application US/08582776C  
; Patent No. 6077510  
; GENERAL INFORMATION:  
; APPLICANT: Lipkin, W. I.

; APPLICANT: Briese, Thomas  
; APPLICANT: Kliche, Stefanie  
; APPLICANT: Schneider, Patrick A.  
; APPLICANT: Stitz, Lothar  
; APPLICANT: Schneemann, Anette  
; TITLE OF INVENTION: Borna Disease Viral Sequences,  
; TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous  
; TITLE OF INVENTION: System Diseases  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski, L.L.P.  
; STREET: 865 South Figueroa Street, 29th Floor  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90017-2576  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: WINDOWS NT  
; SOFTWARE: ASCII DOS TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/582.776C  
; FILING DATE: 04-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/369,822  
; FILING DATE: 06-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/434,831  
; FILING DATE: 04-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Churchill, Margaret A.  
; REGISTRATION NUMBER: 39,944  
; REFERENCE/DOCKET NUMBER: 1279-194C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 213/892-9200  
; TELEFAX: 213/680-4518  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-582-776C-27

Query Match 44.9%; Score 22; DB 3; Length 8;  
Best Local Similarity 66.7%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 TRDEFD 9  
:||||  
Db 1 TADWD 6

Search completed: January 17, 2001, 13:42:57  
Job time: 156 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:42:22 ; Search time 43.61 Seconds  
(without alignments)  
7.057 Million cell updates/sec

Title: US-08-765-837-8  
Sequence: 50  
1 ENMERNCA 9

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 61695

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_36:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	17 R90952	Factor VIII antige
2	25	50.0	7	18 W30412	HRE-I affinity pep
3	24	48.0	9	20 Y48654	Membrane dipeptida
4	24	48.0	9	21 Y66247	HLA-B8-binding HIV
5	23	46.0	7	21 Y61508	Cadherin-7 cell ad
6	23	46.0	7	21 Y62238	Cadherin-14 cell a
7	23	46.0	7	21 Y64309	Cadherin-related n
8	23	46.0	8	10 P93523	Amino acid sequenc
9	23	46.0	8	21 Y61511	Cadherin-7 cell ad
10	23	46.0	8	21 Y62239	Cadherin-14 cell a
11	23	46.0	8	21 Y64312	Cadherin-related n
12	23	46.0	9	17 W49414	Human leucocyte an

13	23	46.0	9	21 Y61514	Cadherin-7 cell ad
14	23	46.0	9	21 Y62234	Cadherin-14 cell a
15	23	46.0	9	21 Y64315	Cadherin-related n
16	22	44.0	6	21 Y64334	Cadherin-related n
17	22	44.0	7	21 Y64337	Cadherin-related n
18	22	44.0	8	18 W35495	Beta2-glycoprotein
19	22	44.0	8	19 W59338	Non-polio enterovi
20	22	44.0	8	20 Y48625	Membrane dipeptida
21	22	44.0	8	20 Y48931	Membrane dipeptida
22	22	44.0	8	21 Y64340	Cadherin-related n
23	22	44.0	8	21 Y64343	Cadherin-related n
24	21	42.0	6	7 Y61430	[1-desmimo-penicill
25	21	42.0	6	21 Y63632	Desmoglein cell ad
26	21	42.0	6	21 Y63933	Desmoglein cell a
27	21	42.0	6	21 Y63945	Desmoglein cell a
28	21	42.0	7	21 Y63635	Desmoglein cell ad
29	21	42.0	7	21 Y63936	Desmoglein cell a
30	21	42.0	7	21 Y63948	Desmoglein cell a
31	21	42.0	8	20 Y48734	Membrane dipeptida
32	21	42.0	8	20 W88221	Pentactinra macro
33	21	42.0	8	21 Y63638	Desmoglein cell ad
34	21	42.0	8	21 Y63939	Desmoglein cell a
35	21	42.0	8	21 Y63951	Desmoglein cell a
36	21	42.0	9	14 R41452	Antigenic peptide
37	21	42.0	9	14 R42988	Murine MHC class I
38	21	42.0	9	15 R51902	Influenza NP 366-3
39	21	42.0	9	15 R55541	Influenza nucleopr
40	21	42.0	9	16 R83943	MHC class I restr
41	21	42.0	9	16 R68025	Influenza nucleopr
42	21	42.0	9	16 R67664	Peptide fragment o
43	21	42.0	9	17 W49679	Human leucocyte an
44	21	42.0	9	17 W49665	Human leucocyte an
45	21	42.0	9	17 W49666	Human leucocyte an

#### ALIGNMENTS

##### RESULT 1

R90952  
ID R90952 standard; peptide; 9 AA.  
AC R90952;  
XX  
XX  
DT 09-SEP-1996 (first entry)  
XX  
DE Factor VIII antigenic peptide corresp. to residues Glu1893-Ala1901.  
XX  
KW Factor VIII; modification; inhibitor activity; blinding; antibody;  
KW von Willebrand factor; immune disorder.  
XX  
OS Synthetic.  
XX  
PN WO9602572-A2.  
XX  
PD 01-FEB-1996.  
XX  
PF 14-JUL-1995; 95WO-BE000068.  
XX  
PR 14-JUL-1994; 94BE-00000666.  
XX  
XX (CROI-) CROIX ROUGE BELGIQUE.  
XX  
PI Di Giambattista M, Laub R;  
XX  
DR WPI; 1996-105861/11.  
XX  
PT Factor VIII antigenic polypeptide fragments and epitope(s) - also  
PT inhibitors of factor VIII and anti-inhibitors, useful for e.g.  
PT preventing and treating immune disorders involving inhibition of  
XX factor VIII binding  
XX  
PS Claim 8; Page 32; 45pp; French.

XX Peptides R90945-64 are derived from the factor VIII protein, esp. from a  
 CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,  
 CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII  
 CC and derived peptides can be used to prevent the activity of inhibitors of  
 CC factor VIII binding to von Willebrand factor, esp. antibodies, thus  
 CC preventing or treating immune disorders.

XX Sequence 9 AA;

Query Match 100.0%; Score 50; DB 17; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENMERNCR 9  
 Db 1 enmerncr 9

RESULT 2  
 W30412  
 ID W30412 standard; peptide; 7 AA.

XX AC W30412;

XX 22-APR-1998 (first entry)

XX HRE-I affinity peptide SEQ ID NO:30.

XX Binding site; HRE-I; screening; H-Ras; target gene; plasmid;  
 KW Inhibition; affinity peptide.

XX Synthetic.

XX W09737030-A1.

XX 09-OCT-1997.

XX 31-MAR-1997; 97WO-JP011105.

XX 03-OCT-1996; 96JP-0263345.

XX 01-APR-1996; 96JP-0101990.

XX (YAWH ) NIPPON STEEL CHEM CO.

PA (YAWH ) NIPPON STEEL CORP.

PI Kato T;

XX WPI; 1997-503115/46.

XX Plasmid for screening peptide(s) which bind target genes - for  
 PT treatment of conditions associated with H-Ras

XX Claim 41; Page 68; 127pp; Japanese.

XX A novel plasmid has been developed for screening peptides which bind  
 CC target genes. The plasmid comprises a sequence encoding the test  
 CC peptide, a promoter comprising the target gene, and a reporter gene  
 CC bound to the promoter. The present sequence represents a specifically  
 CC claimed peptide with affinity to HRE-I. The plasmids are used for  
 CC screening for peptides which bind to target genes. The identified  
 CC peptides can be used for the treatment of conditions associated with  
 CC the inhibition of the expression the genes, and treatment of conditions  
 CC associated with H-Ras.

XX Sequence 7 AA;

Query Match 50.0%; Score 25; DB 18; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RNCR 8  
 Db 4 rncr 7

RESULT 3  
 Y48654  
 ID Y48654 standard; Peptide; 9 AA.

XX AC Y48654;

XX 10-DEC-1999 (first entry)

XX Membrane dipeptidase-binding lung homing peptide #25.

XX Homing peptide; organ; tissue; lung; pancreas; skin; retina; MDP;  
 KW prostate; ovary; lymph node; adrenal gland; liver; gut; tumour;  
 KW membrane dipeptidase.

XX Synthetic.

XX Homo sapiens.

XX W09946284-A2.

XX 16-SEP-1999.

XX 10-MAR-1999; 99WO-US05284.

XX 13-MAR-1998; 98US-0042107.

XX 26-FEB-1999; 99US-0042107.

XX (BURN-) BURNHAM INST.

XX Rajotte D, Pasqualini R, Ruoslahti EI;

XX WPI; 1999-571717/48.

XX New peptides which selectively home to organs or tissues, used for,  
 PT e.g. identifying target ligands and for therapy of pathological  
 PT conditions

XX Example 6; Page 144; 193pp; English.

XX The present invention describes peptides that selectively home to a  
 CC tissue or organ. The peptides can be used for identifying an organ or  
 CC or tissue, for identifying a target molecule expressed by an organ or  
 CC tissue or for treating an organ or tissue pathology, where the organ or  
 CC tissue is selected from prostate, lung, skin, retina, pancreas, gut,  
 CC ovary, adrenal gland, liver, and lymph node. The peptide bind to the  
 CC membrane dipeptidase (MDP). Y48618 to Y49066 represent sequences  
 CC which are used in the exemplification of the present invention.

XX Sequence 9 AA;

Query Match 48.0%; Score 24; DB 20; Length 9;  
 Best Local Similarity 42.9%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENMERNC 7  
 Db 3 qdirnc 9

RESULT 4  
 Y66247  
 ID Y66247 standard; Peptide; 9 AA.

XX AC Y66247;

XX 22-FEB-2000 (first entry)

XX HLA-B8-binding HIV-1 Gag peptide #49.

XX HIV-1; MHC; major histocompatibility complex; Class I; HLA;  
 KW human leukocyte antigen; allele; binding; conserved; genome;  
 KW peptide; targeting; toxic; drug; antibody; antigen; antiviral;  
 KW molecular conjugate therapeutic; diagnosis; treatment; pathogen;  
 KW localisation; quantification; detection; infection; drug resistance;  
 KW immune response.  
 XX Human immunodeficiency virus type 1.  
 OS WO9949893-A1.  
 PN 07-OCT-1999.  
 XX 31-MAR-1999; 99WO-US07111.  
 PF 31-MAR-1998; 98US-0052530.  
 XX (UYBO-) UNIV BOSTON.  
 PA Dellisi C, Berzofsky J, Gulukota K, Vaccaro D, Weng Z, Zhang C;  
 PI WPI; 2000-038361/03.  
 XX Novel methods for designing molecular conjugate therapeutics which are  
 used for diagnosis, imaging and treatment against pathogens -  
 Example 3; Page 46; 62pp; English.  
 XX Y66199-Y66413 are peptides derived from conserved portions of the  
 CC HIV-1 genome which bind to different HLA alleles of MHC (major  
 CC histocompatibility complex) Class I molecules. The peptides are used to  
 CC construct targeting antigens comprising one or more peptides bound to  
 CC the corresponding MHC Class I molecule, which can be used to raise  
 CC antibodies. The antibody may then be used as a targeting vehicle to  
 CC deliver a potentially toxic drug to its target site of action, rather  
 CC than administering it systemically, which may result in adverse side  
 CC effects. The invention relates to improved methods for the design of  
 CC molecular conjugate therapeutics for the diagnosis and treatment of  
 CC infections caused by pathogens with a high mutation rate (such as HIV-1).  
 CC This method involves identifying conserved peptide-encoding regions among  
 CC the genomes of multiple variants of a pathogen, identifying the Class I  
 CC MHC molecules which occur with greatest frequency in a population of  
 CC interest (e.g., human sub-populations), and determining which of the  
 CC peptides bind to the Class I MHC molecules. The MHC-binding peptides and  
 CC the corresponding Class I MHC molecules are selected and used to  
 CC construct targeting antigens, which are in turn used to produce  
 CC targeting antibodies. The methods may be used in localisation,  
 CC quantification and in situ detection of specific peptide-MHC Class I  
 CC complexes and also to detect and treat viral infection. The methods of  
 CC the invention mitigate against the development of viral resistance to  
 CC drugs and to the immune response, as well as providing a solution for  
 CC targeting toxic compounds to destroy viruses sequestered in sites not  
 CC accessible to T cells. In addition, the methods eliminate the virus,  
 CC whereas current therapies only arrest viral replication.  
 XX Sequence 9 AA;

Query Match 48.0%; Score 24; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 NCRA 9  
 Db 1 ncra 4

RESULT 5  
 Y61508  
 ID Y61508 standard; Peptide; 7 AA.  
 XX Y61508;  
 AC

XX 02-MAR-2000 (first entry)  
 DT Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1394.  
 DE Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 XX inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Disulfide-bond 1..7  
 FT WO9957149-A2.  
 XX 11-NOV-1999.  
 PD 05-MAY-1999; 99WO-CA00363.  
 XX 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
 PA Blaschuk OW, Gour BJ, Byers S;  
 PI WPI; 2000-038791/03.  
 XX New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX Claim 36; Page 172; 252pp; English.  
 PS The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioassays. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 CC exemplification of the present invention.  
 XX Sequence 7 AA;

Query Match 46.0%; Score 23; DB 21; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;

Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NMERNC 7  
1:: 11  
Db 2 nidanc 7

RESULT 6  
Y62238  
ID Y62238 standard; Peptide; 7 AA.  
XX  
AC Y62238;  
XX  
DT 02-MAR-2000 (first entry)  
XX  
DE Cadherin-14 cell adhesion recognition cyclic peptide SEQ ID NO:3973.  
XX  
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
KW neurological disease; cyclic.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
FT Disulfide-bond 1..7  
XX  
XX WO9957149-A2.  
XX  
PD 11-NOV-1999.  
XX  
XX 05-MAY-1999; 99WO-CA00363.  
XX  
XX 05-MAY-1998; 98US-0073040.  
PR 06-NOV-1998; 98US-0187859.  
PR 20-JAN-1999; 99US-0234395.  
PR 08-MAR-1999; 99US-0264516.  
XX  
XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
XX  
XX Blaschuk OW, Gour BJ, Byers S;  
XX  
XX WPI; 2000-038791/03.  
XX  
XX New cadherin modulating agents, used for modulating nonclassical  
PT cadherin-mediated functions for treating e.g. cancers, obesity,  
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
PT disease -  
XX  
PS Claim 54; Page 184; 252pp; English.

The present invention describes cadherin modulating agents (NA)  
CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
CC recognition (CAR) sequence. The MAS can be used for modulating  
CC nonclassical cadherin-mediated functions. They can be used for e.g.  
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
CC mammal, enhancing delivery of a drug through the skin of a mammal,  
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
CC expressing cell, preventing or treating obesity in a mammal, stimulating  
CC blood vessel regression in a mammal, enhancing drug delivery to the  
CC central nervous system, treating a demyelinating neurological disease,  
CC increasing vasopermeability in a mammal, enhancing adhesion of  
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a

CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
CC age-related macular degeneration, multiple sclerosis and diabetes. The  
CC products can also be used for detection and diagnosis and in  
CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 7 AA;

Query Match 46.0%; Score 23; DB 21; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMERNC 7  
1:: 11  
Db 2 nidanc 7

RESULT 7  
Y64309  
ID Y64309 standard; Peptide; 7 AA.  
XX  
AC Y64309;  
XX  
DT 02-MAR-2000 (first entry)  
XX  
DE Cadherin-related neuronal receptor CAR cyclic peptide SEQ ID NO:3623.  
XX  
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
KW neurological disease; cyclic.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
FT Disulfide-bond 1..7  
XX  
XX WO9957149-A2.  
XX  
PD 11-NOV-1999.  
XX  
XX 05-MAY-1999; 99WO-CA00363.  
XX  
XX 05-MAY-1998; 98US-0073040.  
PR 06-NOV-1998; 98US-0187859.  
PR 20-JAN-1999; 99US-0234395.  
PR 08-MAR-1999; 99US-0264516.  
XX  
XX (ADHE-) ADHEREX TECHNOLOGIES INC.  
XX  
XX Blaschuk OW, Gour BJ, Byers S;  
XX  
XX WPI; 2000-038791/03.  
XX  
XX New cadherin modulating agents, used for modulating nonclassical  
PT cadherin-mediated functions for treating e.g. cancers, obesity,  
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
PT disease -  
XX  
PS Claim 102; Page 217; 252pp; English.

The present invention describes cadherin modulating agents (NA)  
CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
CC recognition (CAR) sequence. The MAS can be used for modulating  
CC nonclassical cadherin-mediated functions. They can be used for e.g.  
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
CC mammal, enhancing delivery of a drug through the skin of a mammal,  
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
CC expressing cell, preventing or treating obesity in a mammal, stimulating  
CC blood vessel regression in a mammal, enhancing drug delivery to the  
CC central nervous system, treating a demyelinating neurological disease,  
CC increasing vasopermeability in a mammal, enhancing adhesion of  
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a

CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 7 AA;

Query Match 46.0%; Score 23; DB 21; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMERNC 7  
 |::||  
 Db 2 nldsnc 7

RESULT 8  
 P93523  
 ID P93523 standard; protein: 8 AA.  
 AC P93523;  
 XX  
 XX  
 DT 03-JUN-1990 (first entry)  
 XX  
 DE Amino acid sequence of polylinker region of transfer vector pBF133.  
 XX  
 KW Polyhedrin protein; silkworm nuclear polyhedrosis virus; NPV;  
 KW HIV gene; gag; pol; SOR; gp120; gp41; polylinker region; pBF133.  
 XX  
 XX  
 PN JP01074990-A.  
 XX  
 XX 20-MAR-1989.  
 XX  
 PF 17-SEP-1987; 87JP-0231107.  
 XX  
 PR 17-SEP-1987; 87JP-0231107.  
 XX  
 PA (MAED ) MAEDA S.  
 XX  
 XX WPI: 1989-127530/17.  
 DR N-PSDB: N92399.  
 XX

XX New transfer vector pBFs, for HIV diagnosis -  
 PT contg. promoter region of polyhedrin protein gene for slk-worm  
 PT nuclear polyhedrosis virus deoxyribonucleic acid  
 XX  
 PS Fig 2; page 9; 11pp; Japanese.  
 XX  
 CC Transfer vectors pBFs were prepd. contg. promoter region of polyhedrin  
 CC protein gene of silkworm nuclear polyhedrosis virus (NPV) recombined  
 CC with HIV gene gag, pol, SOR, gp120 or gp41 in the polylinker region  
 CC of the vector. Thus, HIV antigenic protein can be produced in large  
 CC amts. and used as an antigen for the diagnosis of HIV or for  
 CC vaccine production.  
 XX  
 SQ Sequence 8 AA;

Query Match 46.0%; Score 23; DB 10; Length 8;  
 Best Local Similarity 62.5%; Pred. No. 2.1e+05;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENMERNCR 8  
 | ||| |  
 Db 1 eqeernsr 8

RESULT 9  
 Y61511  
 ID Y61511 standard; Peptide: 8 AA.  
 XX  
 AC Y61511;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1397.  
 XX

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; Tf-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Disulfide-bond 1..8  
 XX  
 PN WO9957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX

PF 05-MAY-1999; 99WO-CA00363.  
 XX

PR 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX

PA (ADHE-) ADHEREX TECHNOLOGIES INC.

PI Blaschuk OW, Gour BJ, Byers S;

WPI: 2000-038791/03.

XX New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease .  
 XX

PS Claim 36; Page 172; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for

CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 CC exemplification of the present invention.

XX Sequence 8 AA;

Query Match 46.0%; Score 23; DB 21; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMERNC 7  
 I:: II  
 Db 3 nIdanc 8

RESULT 10

Y62239  
 ID Y62239 standard; Peptide: 8 AA.

XX AC Y62239;

XX DT 02-MAR-2000 (first entry)

XX DE Cadherin-14 cell adhesion recognition cyclic peptide SEQ ID NO:3974.

XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; F-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Disulfide-bond 1..8

XX PN WO9957149-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00363.

XX PR 05-MAY-1998; 98US-0073040.

XX PR 06-NOV-1998; 98US-0187859.

XX PR 20-JAN-1999; 99US-0234395.

XX PR 08-MAR-1999; 99US-0264516.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Blaschuk OW, Gour BJ, Byers S;

XX DR WPI; 2000-038791/03.

XX PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -

XX PS Claim 54; Page 184; 252pp; English.

XX CC The present invention describes cadherin modulating agents (MA)

XX CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating

CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 CC exemplification of the present invention.

XX Sequence 8 AA;

Query Match 46.0%; Score 23; DB 21; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMERNC 7  
 I:: II  
 Db 3 nIdanc 8

RESULT 11

Y64312

ID Y64312 standard; Peptide: 8 AA.

XX AC Y64312;

XX DT 02-MAR-2000 (first entry)

XX DE Cadherin-related neuronal receptor CAR cyclic peptide SEQ ID NO:3626.

XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; F-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Disulfide-bond 1..8

XX PN WO9957149-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00363.

XX PR 05-MAY-1998; 98US-0073040.

XX PR 06-NOV-1998; 98US-0187859.

XX PR 20-JAN-1999; 99US-0234395.

XX PR 08-MAR-1999; 99US-0264516.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Blaschuk OW, Gour BJ, Byers S;

XX XX



DR WPI; 2000-038791/03.  
XX New cadherin modulating agents, used for modulating nonclassical  
PT cadherin-mediated functions for treating e.g. cancers, obesity,  
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
PT disease -  
XX  
PS Claim 102; Page 217; 252pp; English.  
XX  
CC The present invention describes cadherin modulating agents (MA)  
CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
CC recognition (CAR) sequence. The MAs can be used for modulating  
CC nonclassical cadherin-mediated functions. They can be used for e.g.  
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
CC mammal, enhancing delivery of a drug through the skin of a mammal,  
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
CC expressing cell, preventing or treating obesity in a mammal, stimulating  
CC blood vessel regression in a mammal, enhancing drug delivery to the  
CC central nervous system, treating a demyelinating neurological disease,  
CC increasing vasopermeability in a mammal, enhancing adhesion of  
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
CC age-related macular degeneration, multiple sclerosis and diabetes. The  
CC products can also be used for detection and diagnosis and in  
CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
CC and Y64573 to Y64643 and 233183 to 233186 represent sequences used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 8 AA;

Query Match 46.0%; Score 23; DB 21; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMNRC 7  
|::||  
Db 3 nidsnc 8

RESULT 12  
W49414  
ID W49414 standard; peptide; 9 AA.  
XX  
AC W49414;  
XX  
DT 05-JUN-1998 (first entry)  
XX  
DE Human leucocyte antigen DQ4 binding peptide #305.  
XX  
KW Human leucocyte antigen; HLA-DQ4; combinatorial library;  
KW autoimmune disease; Chronic articular rheumatism.  
XX  
OS Synthetic.  
XX  
PN JP08151396-A.  
XX  
PD 11-JUN-1996.  
XX  
PF 28-NOV-1994; 94JP-0292657.  
XX  
PR 28-NOV-1994; 94JP-0292657.  
XX  
PA (TEIJ ) TEIJIN LTD.  
XX  
DR WPI; 1996-329479/33.  
XX  
PT HLA-binding oligopeptide and an Immuno::regulator contg it - used in

PT the treatment of auto::Immune disease  
XX  
PS Claim 4; Page 37; 61pp; Japanese.  
XX  
CC This peptide is an example of a peptide which binds to a human leucocyte  
CC antigen HLA-DQ4 molecule. The peptide was isolated from a phagemid  
CC combinatorial library comprising the sequence V05953, by screening with  
CC an HLA-DQ4 molecule. The peptide is used for the treatment of autoimmune  
CC disease, or especially for treatment of viral diseases.  
XX  
SQ Sequence 9 AA;

Query Match 46.0%; Score 23; DB 17; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 RNCRA 9  
|::||  
Db 3 rrcra 7

RESULT 13  
Y61514  
ID Y61514 standard; Peptide; 9 AA.  
XX  
AC Y61514;  
XX  
DT 02-MAR-2000 (first entry)  
XX  
DE Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1400.  
XX  
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
KW neurological disease; cyclic.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..9  
XX  
PN WO9957149-A2.  
XX  
PD 11-NOV-1999.  
XX  
PF 05-MAY-1999; 99WO-CA00363.  
XX  
PR 05-MAY-1998; 98US-0073040.  
PR 06-NOV-1998; 98US-0187859.  
PR 20-JAN-1999; 99US-0234395.  
PR 08-MAR-1999; 99US-0264516.  
XX  
PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
XX  
PI Blaschuk OW, Gour BJ, Byers S;  
XX  
DR WPI; 2000-038791/03.  
XX  
PT New cadherin modulating agents, used for modulating nonclassical  
PT cadherin-mediated functions for treating e.g. cancers, obesity,  
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
PT disease -  
XX  
PS Claim 36; Page 172; 252pp; English.  
XX  
CC The present invention describes cadherin modulating agents (MA)  
CC comprising peptides which comprise a nonclassical cadherin cell adhesion

CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 CC exemplification of the present invention.

XX Sequence 9 AA;

Query Match 46.0%; Score 23; DB 21; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMERNC 7  
 I:: I  
 Db 4 nidanc 9

RESULT 14  
 Y62234

ID Y62234 standard; Peptide: 9 AA.

AC Y62234;

DT 02-MAR-2000 (first entry)

XX Cadherin-14 cell adhesion recognition cyclic peptide SEQ ID NO:1967.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Disulfide-bond 1..9

XX WO9957149-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00363.

XX 05-MAY-1998; 98US-0073040.

PR 06-NOV-1998; 98US-0187859.

PR 20-JAN-1999; 99US-0234395.

PR 08-MAR-1999; 99US-0264516.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

PA Blaschuk OW, Gour BJ, Byers S;

PI

XX

DR WPI; 2000-038791/03.

XX

PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease

XX

PS Claim 54; Page 184; 252pp; English.

XX

CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 CC exemplification of the present invention.

XX Sequence 9 AA;

Query Match 46.0%; Score 23; DB 21; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMERNC 7  
 I:: I  
 Db 4 nidanc 9

RESULT 15

Y64315

ID Y64315 standard; Peptide: 9 AA.

XX Y64315;

XX 02-MAR-2000 (first entry)

XX Cadherin-related neuronal receptor CAR cyclic peptide SEQ ID NO:3629.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Disulfide-bond 1..9

XX WO9957149-A2.

XX

PD 11-NOV-1999.  
 XX  
 PF 05-MAY-1999; 99WO-CA00363.  
 XX  
 PR 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 XX WPI; 2000-038791/03.  
 DR  
 XX  
 XX New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX  
 PS Claim 102; Page 217; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin-expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,  
 CC age-related macular degeneration, multiple sclerosis and diabetes. The  
 CC products can also be used for detection and diagnosis and in  
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,  
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 46.0%; Score 23; DB 21; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMERNC 7  
 I:: I  
 Db 4 nidsnc 9

Search completed: January 17, 2001, 13:42:23  
 Job time: 137 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:43:37 ; Search time 36.59 Seconds  
(without alignments)  
16.701 Million cell updates/sec

Title: US-08-765-837-8

Perfect score: 50

Sequence: 1 ENMERNCA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 787

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66:\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	40.0	9	B28495	conopressin S - co
2	20	40.0	9	A28495	conopressin G - co
3	20	40.0	9	PC2021	oxytocin-related p
4	20	40.0	9	A39040	lysine-conopressin
5	17	34.0	8	I57532	gene Tn1slow prote
6	16	32.0	6	I67345	MHC H2-K-k cell su
7	16	32.0	9	A14666	oxytocin - hippopo
8	16	32.0	9	A92774	oxytocin - spotted
9	16	32.0	9	A93147	oxytocin - finback
10	16	32.0	9	A93408	oxytocin - Austral
11	16	32.0	9	B90667	oxytocin - rabbit
12	16	32.0	9	S06375	arginine vasotocin
13	16	32.0	9	B61364	vasotocin - common
14	15	30.0	6	A24952	acylaminoacyl-pept
15	15	30.0	7	S38516	mablinin II chain
16	15	30.0	8	A25836	L-serine dehydrata
17	15	30.0	8	A54823	olfactory receptor
18	15	30.0	8	I57018	gene cfr protein
19	15	30.0	8	B54823	olfactory receptor
20	15	30.0	9	A61364	isotocin - common
21	15	30.0	9	A29477	diuretic neuropept
22	14	28.0	8	S59622	metallothionein is
23	14	28.0	8	A37521	R-phycoerythrin ga
24	14	28.0	8	A39892	P element, P cytot
25	14	28.0	9	B45020	probable minipolyp
26	14	28.0	9	PH0921	T-cell receptor be
27	13	26.0	6	B44510	hypothetical prote
28	13	26.0	6	PC4127	hypothetical 6 pro
29	13	26.0	9	S66608	quinoline 2-oxidor

#### ALIGNMENTS

##### RESULT 1

B28495  
conopressin S - cone shell (Conus striatus)  
N:Alternate names: Arg-vasopressin-S  
C:Species: Conus striatus (striated cone)  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 25-Apr-1997  
C:Accession: B28495  
R:Cruz, L.J.; de Santos, V.; Zafaralla, G.C.; Ramilo, C.A.; Zelkus, R.; Gray, W.R.; C  
J. Biol. Chem. 262, 15821-15824, 1987  
A:Title: Invertebrate vasopressin/oxytocin homologs. Characterization of peptides frc  
A:Reference number: A92617; MUID:88058932  
A:Accession: B28495  
A:Molecule type: protein  
A:Residues: 1-9 <CRU>  
C:Superfamily: oxytocin-neurophysin  
C:Keywords: amidated carboxyl end; venom  
F:1-6/Disulfide bonds: #status experimental  
F:9/Modified site: amidated carboxyl end (Gly) #status experimental

30	13	26.0	9	2	A26363	cardioactive pepti
31	13	26.0	9	2	S27233	cardioactive pepti
32	13	26.0	9	2	S39767	cardioactive pepti
33	12	24.0	4	2	E44823	synaptosomal-assoc
34	12	24.0	5	2	B31836	20K protein - Rick
35	12	24.0	5	2	S65726	hemoglobin, extrac
36	12	24.0	6	2	I49424	cytotoxic T-lympho
37	12	24.0	7	1	XEVDGD	galactose oxidase
38	12	24.0	7	2	A12016	formylglycinamide
39	12	24.0	8	2	S43971	tumor-associated a
40	12	24.0	8	2	S43972	tumor-associated a
41	12	24.0	8	2	PC1002	leucine--tRNA liga
42	12	24.0	8	2	PH0803	T-cell receptor al
43	12	24.0	9	2	S19329	sperm-activating p
44	12	24.0	9	2	I54379	gene NF2 protein -
45	12	24.0	9	2	C60070	gastrin - domestic

Query Match 40.0%; Score 20; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 RNC 7  
DB 4 RNC 6

##### RESULT 2

A28495  
conopressin G - cone shell (Conus geographus)  
N:Alternate names: Lys-conopressin-G  
C:Species: Conus geographus (geography cone)  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 25-Apr-1997  
C:Accession: A28495  
R:Cruz, L.J.; de Santos, V.; Zafaralla, G.C.; Ramilo, C.A.; Zelkus, R.; Gray, W.R.; O  
J. Biol. Chem. 262, 15821-15824, 1987  
A:Title: Invertebrate vasopressin/oxytocin homologs. Characterization of peptides frc  
A:Reference number: A92617; MUID:88058932  
A:Accession: A28495  
A:Molecule type: protein  
A:Residues: 1-9 <CRU>  
C:Superfamily: oxytocin-neurophysin  
C:Keywords: amidated carboxyl end; venom  
F:1-6/Disulfide bonds: #status experimental  
F:9/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 40.0%; Score 20; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



RESULT 8  
A92774  
oxytocin - spotted ratfish  
N:Alternate names: neurophysin I  
N:Contains: neurophysin 1; oxytocin  
C:Species: Hydrolagus collii (spotted ratfish)  
C:Date: 30-Oct-1992 #sequence\_revision 30-Oct-1992 #text\_change 20-Mar-1998  
C:Accession: A92774; A01450; B01450  
R:Pickering, B.T.; Heller, H.  
J. Endocrinol. 45, 597-606, 1969  
A:Title: Oxytocin as a neurohypophyseal hormone in the holocephalian elasmobranch fish,  
A:Reference number: A92774; MUID:70088110  
A:Accession: A92774  
A:Molecule type: protein  
A:Residues: 1-9 <PIC>  
C:Comment: Oxytocin is followed by neurophysin 1 in the precursor.  
C:Superfamily: oxytocin-neurophysin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus  
F:1-6/Disulfide bonds: #status predicted  
F:9/Modified site: amidated carboxyl end (Gly) #status predicted

Query Match 32.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RNC 7  
:||  
Db 4 QNC 6

RESULT 9  
A93147  
oxytocin - finback whale  
N:Alternate names: neurophysin I  
N:Contains: neurophysin 1; oxytocin  
C:Species: Balaeoptera physalus (finback whale, common rorqual)  
C:Date: 30-Oct-1992 #sequence\_revision 30-Oct-1992 #text\_change 20-Mar-1998  
C:Accession: A93147; A01450; B01450  
R:Acher, R.; Chauvet, J.; Chauvet, M.T.  
Nature 201, 191-192, 1964  
A:Title: Isolation of finback whale oxytocin and vasopressin.  
A:Reference number: A93147  
A:Accession: A93147  
A:Molecule type: protein  
A:Residues: 1-9 <ACH>  
C:Comment: Oxytocin is followed by neurophysin 1 in the precursor.  
C:Superfamily: oxytocin-neurophysin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus  
F:1-6/Disulfide bonds: #status experimental  
F:9/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 32.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RNC 7  
:||  
Db 4 QNC 6

RESULT 10  
A93408  
oxytocin - Australian echidna  
N:Alternate names: neurophysin I  
N:Contains: neurophysin 1; oxytocin  
C:Species: Tachyglossus aculeatus (Australian echidna)  
C:Date: 30-Oct-1992 #sequence\_revision 30-Oct-1992 #text\_change 20-Mar-1998  
C:Accession: A93408; A01450; B01450  
R:Acher, R.; Chauvet, J.; Chauvet, M.T.  
Nature New Biol. 244, 124-126, 1973  
A:Title: Neurohypophyseal hormones and evolution of tetrapods.

Query Match 32.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

A:Reference number: A93408; MUID:73223515  
A:Accession: A93408  
A:Molecule type: protein  
A:Residues: 1-9 <ACH>  
C:Comment: Oxytocin is followed by neurophysin 1 in the precursor.  
C:Superfamily: oxytocin-neurophysin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus  
F:1-6/Disulfide bonds: #status predicted  
F:9/Modified site: amidated carboxyl end (Gly) #status predicted

Query Match 32.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RNC 7  
:||  
Db 4 QNC 6

RESULT 11  
B90667  
oxytocin - rabbit  
N:Alternate names: neurophysin I  
N:Contains: neurophysin 1; oxytocin  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 30-Oct-1992 #sequence\_revision 30-Oct-1992 #text\_change 20-Mar-1998  
C:Accession: B90667; A01450; B01450  
R:Chauvet, J.; Chauvet, M.T.; Acher, R.  
Biochimie 53, 1099-1104, 1971  
A:Title: Evolution des hormones neurohypophysiales: isolement des principes actifs du  
A:Reference number: A90667; MUID:72215060  
A:Accession: B90667  
A:Molecule type: protein  
A:Residues: 1-9 <CHA>  
C:Comment: Oxytocin is followed by neurophysin 1 in the precursor.  
C:Superfamily: oxytocin-neurophysin  
C:Keywords: amidated carboxyl end; hormone; hypothalamus  
F:1-6/Disulfide bonds: #status experimental  
F:9/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 32.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RNC 7  
:||  
Db 4 QNC 6

RESULT 12  
S06375  
arginine vasotocin - sea lamprey  
C:Species: Petromyzon marinus (sea lamprey)  
C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 31-Mar-1997  
C:Accession: S06375  
R:Lane, T.F.; Sower, S.A.; Kawauchi, H.  
Gen. Comp. Endocrinol. 70, 152-157, 1988  
A:Title: Arginine vasotocin from the pituitary gland of the lamprey (Petromyzon marinus)  
A:Reference number: S06375; MUID:88225976  
A:Accession: S06375  
A:Molecule type: protein  
A:Residues: 1-9 <LAN>  
C:Keywords: amidated carboxyl end; hormone; neuropeptide; pituitary  
F:1-6/Disulfide bonds: #status experimental  
F:9/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 32.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RNC 7  
:||  
Db 4 QNC 6

## RESULT 13

B61364  
vasotocin - common carp  
C:Species: Cyprinus carpio (common carp)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 20-Jun-2000  
C:Accession: B61364  
R:Acher, R.; Chauvet, J.; Chauvet, M.T.; Crepy, D.  
Comp. Biochem. Physiol. A 14, 245-254, 1965  
A:Title: Caracterisation des hormones neurohypophysaires d'un poisson osseux d'eau douce  
A:Reference number: A61364  
A:Accession: B61364  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <ACH>  
C:Keywords: amidated carboxyl end; neuropeptide; posterior pituitary

Query Match 32.0%; Score 16; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 RNC 7  
:||  
Db 4 QNC 6

## RESULT 14

A49792  
acylaminoacyl-peptidase (EC 3.4.19.1) - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999  
C:Accession: A49792  
R:Krishna, R.G.; Chin, C.C.Q.; Wold, F.  
Anal. Biochem. 199, 45-50, 1991  
A:Title: N-terminal sequence analysis of N(alpha)-acetylated proteins after unblocking  
A:Reference number: A49792; MUID:92222120  
A:Accession: A49792  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-6 <KRI>  
C:Keywords: acetylated amino end; hydrolase; omega peptidase  
F:1/Modified site: acetylated amino end (Met) #status experimental

Query Match 30.0%; Score 15; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MER 5  
|||  
Db 1 MER 3

## RESULT 15

S38516  
mablinin II chain A - Yunnan caper (fragments)  
C:Species: Capparis masakai (Yunnan caper)  
C:Date: 08-Jun-1994 #sequence\_revision 27-Feb-1997 #text\_change 13-Mar-1997  
C:Accession: S38516  
R:Nirasawa, S.; Liu, X.; Nishino, T.; Kurihara, Y.  
Biochim. Biophys. Acta 1202, 277-280, 1993  
A:Title: Disulfide bridge structure of the heat-stable sweet protein mablinin II.  
A:Reference number: S38516; MUID:94002261  
A:Accession: S38516  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <NIR>

Query Match 30.0%; Score 15; DB 2; Length 7;  
Best Local Similarity 40.0%; Pred. No. 2e+05;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 ERNCR 8  
:| |:  
Db 2 QRACQ 6

Search completed: January 17, 2001, 13:43:38  
Job time: 157 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:47:10 ; Search time 20.9 seconds  
(without alignments)  
13.907 Million cell updates/sec

Title: US-08-765-837-8  
Perfect score: 50  
Sequence: 1 ENMERNCR 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 213

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	40.0	9	1 CONO_CONGE	P05486 conus geogr
2	20	40.0	9	1 CONO_CONST	P05487 conus stria
3	20	40.0	9	1 OXYT_EISFO	P42998 eisenia foe
4	20	40.0	9	1 OXYT_OCTVU	P80027 octopus vul
5	16	32.0	9	1 OXYT_CYPCA	P23879 cyprinus ca
6	16	32.0	9	1 OXYT_RABIT	P32878 oryctolagus
7	16	32.0	9	1 OXYV_SQUAC	P43000 squalus aca
8	15	30.0	6	1 ACPH_RABIT	P25154 oryctolagus
9	15	30.0	9	1 DNFL_LOCFI	P16339 locusta mig
10	15	30.0	9	1 ISOT_CYPCA	P42993 cyprinus ca
11	15	30.0	9	1 OXYA_SCYCA	P42996 scyllorhinu
12	15	30.0	9	1 OXYA_SQUAC	P42998 squalus aca
13	15	30.0	9	1 OXYF_SCYCA	P42997 scyllorhinu
14	15	30.0	9	1 OXYT_RAJCL	P42994 raja clavav
15	15	30.0	9	1 ULAE_HUMAN	P31931 homo sapien
16	14	28.0	9	1 RS11_SALTY	O54296 salmonella
17	13	26.0	7	1 FAR2_PROCL	P38498 procambarus
18	13	26.0	8	1 FAR3_HOMAM	P41486 homarus ame
19	13	26.0	9	1 CCAP_CARMA	P38556 carcinus ma
20	12	24.0	7	1 IGAO_DACDE	P06294 dactylium d
21	12	24.0	8	1 B4AK_PORGI	P81886 porphyromon
22	12	24.0	9	1 SAP_STOVA	P24047 stomopneute
23	11	22.0	6	1 VP19_HSV1K	P23210 herpes simp
24	11	22.0	7	1 FAR1_PROCL	P38499 procambarus
25	11	22.0	7	1 FAR4_CALVO	P41487 calliphora
26	11	22.0	8	1 FAR4_HOMAM	P81010 fusarium so
27	11	22.0	8	1 FUSS_FUSSO	P30087 homarus ame
28	11	22.0	8	1 UPAL_HUMAN	P30087 homo sapien
29	11	22.0	9	1 FIBB_PAPAN	P19344 papio anubi
30	11	22.0	9	1 FIBB_PAPHA	P19343 papio hamad
31	11	22.0	9	1 FIBB_THEGE	P19342 theopithec
32	11	22.0	9	1 FLA2_TREHY	P80159 treponema h
33	11	22.0	9	1 NEUU_CAVPO	P34966 cavia porce

34 11 22.0 9 1 OXYT\_BUFRE p42995 bufo regula  
35 10 20.0 5 1 BIOB\_SALTY p12678 salmonella  
36 10 20.0 8 1 HTF1\_PERAM p04548 periplaneta  
37 10 20.0 8 1 HTF1\_TENMO p25419 tenebrio mo  
38 10 20.0 8 1 NS3\_MYCTU p81152 mycobacteri  
39 10 20.0 8 1 UH09\_RAT p56575 rattus norv  
40 10 20.0 9 1 FABF\_RHIME p56502 rhizobium m  
41 10 20.0 9 1 TAL1\_PICJA p17440 pichia jadi  
42 10 20.0 9 1 TAL3\_PICJA p17441 pichia jadi  
43 10 20.0 9 1 UF02\_MOUSE p38640 mus musculu  
44 10 20.0 9 1 ULAH\_HUMAN p31934 homo sapien  
45 10 20.0 9 1 ULAK\_MOUSE p99031 mus musculu

## ALIGNMENTS

RESULT 1  
CONO\_CONGE STANDARD; PRT; 9 AA.  
AC P05486;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE LYS-CONOPRESSIN G.  
OS Conus geographus (Geography cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=88058932; PubMed=3680228;  
RA Cruz L.J., de Santos V., Zafaralla G.C., Ramilo C.A., Zeilku R.D.,  
RA Gray W.R., Olivera B.M.;  
RT "Invertebrate vasopressin/oxytocin homologs. Characterization of  
peptides from Conus geographus and Conus straitus venoms.";  
RL J. Biol. Chem. 262:15821-15824(1987).  
RN [2]  
RP REVIEW.  
RX MEDLINE=89024586; PubMed=3052286;  
RA Gray W.R., Olivera B.M., Cruz L.J.;  
RT "Peptide toxins from venomous Conus snails.";  
RL Annu. Rev. Biochem. 57:665-700(1988).  
CC -I- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
DR PIR: A28495; A28495.  
DR INTERPRO: IPR000981; -;  
DR PFAM: PF00220; hormone4; 1.  
DR PROSITE: PS00264; NEUROHYPOPHYS\_HORM; 1.  
KW Hormone; Amidation.  
FT DISULFID 1 6  
FT MOD\_RES 9 9  
SQ SEQUENCE 9 AA; 1037 MW; D4FC276EB4540059 CRC64;  
AMIDATION.

Query Match 40.0%; Score 20; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 5 RNC 7  
|||  
Db 4 RNC 6

RESULT 2  
CONO\_CONST STANDARD; PRT; 9 AA.  
AC P05487;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE ARG-CONOPRESSIN S.  
OS Conus striatus (Striated cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.



CC -1- FUNCTION: ANTIDIURETIC HORMONE.  
 CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR PIR: B61364; B61364.  
 DR PIR: S06375; S06375.  
 DR INTERPRO: IPR000981; .  
 DR PFAM: PF00220; hormone4; 1.  
 DR PROSITE: PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Amidation.  
 FT DISULFID 1 6  
 FT MOD\_RES 9 9  
 SQ SEQUENCE 9 AA; 1053 MW; 17EB176EB456D04B CRC64;

Query Match 32.0%; Score 16; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 RNC 7  
 Db 4 QNC 6

## RESULT 6

ID OXYT\_RABIT STANDARD; PRT; 9 AA.  
 AC P32878; P01188;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE OXYTOCIN (OXYTOCIN).  
 OS Oryctolagus cuniculus (Rabbit), Hippopotamus amphibius (Hippopotamus),  
 OS Balaenoptera physalus (Finback whale) (Common rorqual),  
 OS Tachyglossus aculeatus aculeatus (Australian echidna), and  
 OS Hydroloagus collei (Spotted ratfish) (Pacific ratfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=RABBIT;  
 RX MEDLINE=72215060; PubMed=5150741;  
 RA Chauvet J., Chauvet M.-T., Acher R.;  
 RT "Evolution of neurohypophyseal hormones: isolation of active  
 RT principles from rabbits and rats.";  
 RL Biochimie 53:1099-1104(1971).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=H.AMPHIBIUS;  
 RX MEDLINE=71232719; PubMed=5406007;  
 RA Ferguson D.R., Pickering B.T.;  
 RT "Arginine and lysine vasopressins in the hippopotamus  
 RT neurohypophysis.";  
 RL Gen. Comp. Endocrinol. 13:425-429(1969).  
 RN [3]  
 RP SEQUENCE.  
 RC SPECIES=B.PHYSALUS;  
 RA Acher R., Chauvet J., Chauvet M.-T.;  
 RT "Isolation of finback whale oxytocin and vasopressin.";  
 RL Nature 201:191-192(1964).  
 RN [4]  
 RP SEQUENCE.  
 RC SPECIES=A.ACULEATUS;  
 RX MEDLINE=73233515; PubMed=4515919;  
 RA Acher R., Chauvet J., Chauvet M.-T.;  
 RT "Neurohypophyseal hormones and evolution of tetrapods.";  
 RL Nature New Biol. 244:124-126(1973).  
 RN [5]  
 RP SEQUENCE.  
 RC SPECIES=H.COLLEI;  
 RX MEDLINE=70088110; PubMed=5366118;  
 RA Pickering B.T., Heller H.;  
 RT "Oxytocin as a neurohypophyseal hormone in the holocephalian  
 RT elasmobranch fish, Hydroloagus collei.";  
 RL J. Endocrinol. 45:597-606(1969).

CC -1- FUNCTION: OXYTOCIN CAUSES CONTRACTION OF THE SMOOTH MUSCLE OF THE  
 CC UTERUS AND OF THE MAMMARY GLAND.  
 CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR PIR: A91466; A91466.  
 DR PIR: A92774; A92774.  
 DR PIR: A93147; A93147.  
 DR PIR: A93408; A93408.  
 DR PIR: B90667; B90667.  
 DR PDB: 1XY1; 15-OCT-90.  
 DR PDB: 1XY2; 15-OCT-90.  
 DR INTERPRO: IPR000981; .  
 DR PFAM: PF00220; hormone4; 1.  
 DR PROSITE: PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Hypothalamus; Amidation; 3D-structure.  
 FT DISULFID 1 6  
 FT MOD\_RES 9 9  
 SQ SEQUENCE 9 AA; 1010 MW; 17F8376EB456D04B CRC64;

Query Match 32.0%; Score 16; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 RNC 7  
 Db 4 QNC 6

## RESULT 7

ID OXYV\_SQUAC STANDARD; PRT; 9 AA.  
 AC P43000;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE VALITOCIN.  
 OS Squalus acanthias (Spiny dogfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squala; Squaloidae; Squalidae; Squalus.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=730311727; PubMed=5083097;  
 RA Acher R., Chauvet J., Chauvet M.-T.;  
 RT "Phylogeny of the neurohypophyseal hormones. Two new active peptides  
 RT isolated from a cartilaginous fish, Squalus acanthias.";  
 RL Eur. J. Biochem. 29:12-19(1972).  
 RN [2]  
 RP SEQUENCE.  
 RX MEDLINE=72128038; PubMed=4622083;  
 RA Acher R., Chauvet J., Chauvet M.-T., Fontaine M.;  
 RT "Identification of 2 new neurohypophyseal hormones, valitocin (Val8-  
 RT oxytocin) and aspartocin (Asn4-oxytocin) in a selachian fish, the  
 RT spiny dog-fish (Squalus acanthias).";  
 RL C. R. Acad. Sci., D, Sci. Nat. 274:313-316(1972).  
 CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR INTERPRO: IPR000981; .  
 DR PFAM: PF00220; hormone4; 1.  
 DR PROSITE: PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Amidation.  
 FT DISULFID 1 6  
 FT MOD\_RES 9 9  
 SQ SEQUENCE 9 AA; 996 MW; 17EDD76EB456D04B CRC64;

Query Match 32.0%; Score 16; DB 1; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 RNC 7  
 Db 4 QNC 6

```

RESULT 8
ACPH_RABBIT
ID ACPH_RABBIT STANDARD; PRT; 6 AA.
AC P25154;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ACYLAMINO-ACID-RELEASING ENZYME (EC 3.4.19.1) (ACYL-PEPTIDE HYDROLASE)
DE (APH) (ACYLAMINOACYL-PEPTIDASE) (FRAGMENT).
GN APEH.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE.
RC TISSUE=MUSCLE;
RX MEDLINE=92222120; PubMed=1807161;
RA Krishna R.G., Chin C.C.Q., Wold F.;
RT "N-terminal sequence analysis of N alpha-acetylated proteins after
unblocking with N-acylaminoacyl-peptide hydrolase.";
RL Anal. Biochem. 199;45-50(1991).
CC -|- FUNCTION: THIS ENZYME CATALYZES THE HYDROLYSIS OF THE AMINO-
TERMINAL PEPTIDE BOND OF AN N-ACETYLATED PEPTIDE TO GENERATE
AN N-ACETYLATED AA AND A PEPTIDE WITH A FREE AMINO-TERMINUS.
CC IT PREFERENTIALLY CLEAVES OFF AC-ALA, AC-MET AND AC-SER.
CC -|- CATALYTIC ACTIVITY: ACYLAMINOACYL-PEPTIDE + H(2)O = ACYLAMINO
ACID + PEPTIDE.
CC -|- SUBUNIT: HOMOTETRAMER.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9C; ALSO KNOWN AS THE
PROLYL OLIGOPEPTIDASE FAMILY.
CC MEROPS: S09.004; -.
DR INTERPRO: IPR002471; -.
DR PROSITE: PS00708; PRO-ENDOPEP_SER; PARTIAL.
KW Hydrolase; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT NON_TER 6 6
FT SEQUENCE 6 AA; 775 MW; 6732DC640B16F000 CRC64;

Query Match 30.0%; Score 15; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MER 5
Db 1 MER 3

RESULT 9
DNF1_LOCF1
ID DNF1_LOCF1 STANDARD; PRT; 9 AA.
AC P16339;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE LOCUPRESSIN (DIURETIC NEUROPEPTIDE F1/F2).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
RN [1]
RP SEQUENCE.
RC TISSUE-SUBOESOPHAGEAL GANGLION, AND THORACIC GANGLION;
RX MEDLINE=88077077; PubMed=3689410;
RA Proux J.P., Miller C.A., Li J.P., Carney R.L., Girardie A.,
RA Delage M., Schooley D.A.;
RT "Identification of an arginine vasopressin-like diuretic hormone from
Locusta migratoria.";
RL Biochem. Biophys. Res. Commun. 149:180-186(1987).
CC -|- FUNCTION: DIURETIC HORMONE.
CC -|- SUBUNIT: F2 IS AN ANTIPARALLEL DISULFIDE LINKED DIMER OF F1.
CC -|- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
PIR; A29477; A29477.

```

```

DR INTERPRO: IPR000981; -.
DR PFAM: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Neuropeptide; Amidation.
FT DISULFID 1 6 IN F1.
FT DISULFID 1 1 INTERCHAIN (WITH C-6') (IN F2).
FT DISULFID 6 6 INTERCHAIN (WITH C-1') (IN F2).
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 976 MW; 56EB176EB451A057 CRC64;

Query Match 30.0%; Score 15; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 NC 7
Db 5 NC 6

RESULT 10
ISOT_CYPCA
ID ISOT_CYPCA STANDARD; PRT; 9 AA.
AC P42993;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE ISOTOCIN.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.
RN [1]
RP SEQUENCE.
RC TISSUE-PITUITARY;
RA Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT "Characterization of neurohypophyseal hormones from a fresh water bony
fish, the carp (Cyprinus carpio). Comparison with hormones from sea
water bony fishes.";
RL Comp. Biochem. Physiol. 14:245-254(1965).
CC -|- FUNCTION: ANTIDIURETIC HORMONE.
CC -|- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
PIR; A61364; A61364.
DR INTERPRO: IPR000981; -.
DR PFAM: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9
FT SEQUENCE 9 AA; 969 MW; 17FF476EB455B04B CRC64;

Query Match 30.0%; Score 15; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 NC 7
Db 5 NC 6

RESULT 11
OXVA_SCYCA
ID OXVA_SCYCA STANDARD; PRT; 9 AA.
AC P42956;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE ASVATOCIN.
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.

```

```

RN [1]
RP SEQUENCE.
RC TISSUE=PIUITARY;
RX MEDLINE=95062247; PubMed=7972045;
RA Chauvet J., Rouille Y., Chauveau C., Chauvet M.-T., Acher R.;
RT "Special evolution of neurohypophysial hormones in cartilaginous
RT fishes: asvatocin and phasvatocin, two oxytocin-like peptides
RT isolated from the spotted dogfish (Scyliorhinus caniculus).";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11266-11270(1994).
CC -1- FUNCTION: DISPLAYS OXYTOCIC ACTIVITY ON RAT UTERUS.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR INTERPRO: IPR000981; -.
DR PFAM: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 982 MW; 17EDD76EB44404B CRC64;

Query Match 30.0%; Score 15; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 NC 7
DB 5 NC 6

RESULT 12
ID OXYA_SQUAC STANDARD; PRT; 9 AA.
AC P42999;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE ASPARTOCIN (ASPARTOCIN).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
RN [1]
RP SEQUENCE.
RX MEDLINE=73031727; PubMed=5083097;
RA Acher R., Chauvet J., Chauvet M.-T.;
RT "Phylogeny of the neurohypophysial hormones. Two new active peptides
RT isolated from a cartilaginous fish, Squalus acanthias.";
RL Eur. J. Biochem. 29:12-19(1972).
RN [2]
RP SEQUENCE.
RX MEDLINE=72128038; PubMed=4622083;
RA Acher R., Chauvet J., Chauvet M.-T., Fontaine M.;
RT "Identification of 2 new neurohypophysial hormones, valitocin (Val8-
RT oxytocin) and aspartocin (Asn4-oxytocin) in a selachian fish, the
RT spiny dog-fish (Squalus acanthias).";
RL C. R. Acad. Sci., D, Sci. Nat. 274:313-316(1972).
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR INTERPRO: IPR000981; -.
DR PFAM: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 996 MW; 17F8376EB44404B CRC64;

Query Match 30.0%; Score 15; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 NC 7
DB 5 NC 6

```

```

RESULT 13
OXYF_SCYCA STANDARD; PRT; 9 AA.
AC P42997;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PHASVATOCIN.
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
RN [1]
RP SEQUENCE.
RC TISSUE=PIUITARY;
RX MEDLINE=95062247; PubMed=7972045;
RA Chauvet J., Rouille Y., Chauveau C., Chauvet M.-T., Acher R.;
RT "Special evolution of neurohypophysial hormones in cartilaginous
RT fishes: asvatocin and phasvatocin, two oxytocin-like peptides
RT isolated from the spotted dogfish (Scyliorhinus caniculus).";
RL Proc. Natl. Acad. Sci. U.S.A. 91:11266-11270(1994).
CC -1- FUNCTION: DISPLAYS OXYTOCIC ACTIVITY ON RAT UTERUS.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR INTERPRO: IPR000981; -.
DR PFAM: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 1016 MW; 17EDD76EB4449DB CRC64;

Query Match 30.0%; Score 15; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 8.9e+04;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 NC 7
DB 5 NC 6

RESULT 14
OXYT_RAJCL STANDARD; PRT; 9 AA.
AC P42994;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUMITOCIN.
OS Raja clavata (Thornback ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosqualea; Pristiogalea; Batoidea;
OC Rajiformes; Rajidae; Raja.
RN [1]
RP SEQUENCE.
RX MEDLINE=66123415; PubMed=5880565;
RA Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT "Phylogeny of neurohypophysal peptides: isolation of a new hormone,
RT glumitocin (Ser 4-Gln 8-oxytocin) present in a cartilaginous fish,
RT the ray (Raja clavata).";
RL Blochim. Biophys. Acta 107:393-396(1965).
CC -1- FUNCTION: ANTIDIURETIC HORMONE.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR INTERPRO: IPR000981; -.
DR PFAM: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 984 MW; 17E9C76EB455B04B CRC64;

```

Query Match 30.0%; Score 15; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 NC 7  
DB 5 NC 6

RESULT 15  
ULAE\_HUMAN  
ID ULAE\_HUMAN STANDARD; PRT; 9 AA.  
AC P31931;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF LIVER TISSUE (SPOT 115) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE.  
RC TISSUE=LIVER;  
RX MEDLINE=94147969; PubMed=8313870;  
RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,  
RA Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;  
RT "Human liver protein map: update 1993.";  
RL Electrophoresis 14:1216-1222(1993).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 5.5, ITS MW IS: 11 KDA.  
DR SWISS-2DPAGE; P31931; HUMAN.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1096 MW; 477B3B173AE729C7 CRC64;

Query Match 30.0%; Score 15; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MERN 6  
DB 6 LEKN 9

Search completed: January 17, 2001, 13:47:10  
Job time: 279 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:44:41 ; Search time 58.21 Seconds  
(without alignments)  
18.122 Million cell updates/sec

Title: US-08-765-837-8  
Perfect score: 50  
Sequence: 1 ENMERNCA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 467

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_invertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	42.0	8	Q9Y4J3	Q9Y4J3 homo sapien
2	18	36.0	8	Q9Y4J4	Q9Y4J4 homo sapien
3	16	32.0	9	Q48686	Q48686 lactococcus
4	15	30.0	8	Q9SB24	Q9SB24 nicotiana t
5	15	30.0	8	Q85562	Q85562 moloney mur
6	15	30.0	9	Q16386	Q16386 homo sapien
7	15	30.0	9	Q16276	Q16276 homo sapien
8	15	30.0	9	Q57605	Q57605 squash leaf
9	15	30.0	9	Q57606	Q57606 squash leaf
10	14	28.0	7	Q9YVE3	Q9YVE3 human adeno
11	14	28.0	7	Q9YI90	Q9YI90 human adeno
12	14	28.0	7	Q9YI90	Q9YI90 human adeno
13	14	28.0	9	Q31363	Q31363 borrelia ga
14	14	28.0	9	Q9R792	Q9R792 borrelia bu
15	14	28.0	9	Q9R3T0	Q9R3T0 borrelia af
16	14	28.0	9	Q12096	Q12096 caprine art
17	14	28.0	9	Q12098	Q12098 caprine art
18	14	28.0	9	Q12100	Q12100 caprine art
19	14	28.0	9	Q12102	Q12102 caprine art

20	14	28.0	9	12	Q12104	Q12104 caprine art
21	13	26.0	8	2	Q9R3X0	Q9R3X0 planktothri
22	13	26.0	9	2	P82568	P82568 streptococc
23	13	26.0	9	6	Q9TRSO	Q9TRSO oryctolagus
24	12	24.0	8	2	Q45889	Q45889 clostridium
25	12	24.0	8	4	Q15900	Q15900 homo sapien
26	12	24.0	8	4	Q16468	Q16468 homo sapien
27	12	24.0	8	4	Q9P285	Q9P285 homo sapien
28	12	24.0	8	8	Q19957	Q19957 gossypium h
29	12	24.0	8	8	Q36898	Q36898 nicotiana p
30	12	24.0	9	3	Q9UR18	Q9UR18 sclerotium
31	12	24.0	9	4	Q15892	Q15892 homo sapien
32	12	24.0	9	4	Q15896	Q15896 homo sapien
33	12	24.0	9	8	Q31653	Q31653 anser caeru
34	12	24.0	9	12	Q69473	Q69473 human herpe
35	11	22.0	8	2	Q32560	Q32560 escherichia
36	11	22.0	8	2	Q9S6D5	Q9S6D5 escherichia
37	11	22.0	8	4	Q9Y4X6	Q9Y4X6 homo sapien
38	11	22.0	8	4	Q9UDZ4	Q9UDZ4 homo sapien
39	11	22.0	8	8	Q9TD02	Q9TD02 terranatos
40	11	22.0	8	9	Q37854	Q37854 bacterioph
41	11	22.0	8	10	Q40659	Q40659 oryza sativ
42	11	22.0	8	11	Q62933	Q62933 rattus norv
43	11	22.0	8	12	Q84156	Q84156 orf virus
44	11	22.0	8	13	P79940	P79940 xenopus lae
45	11	22.0	9	2	Q57328	Q57328 aeromonas s

#### ALIGNMENTS

RESULT 1  
Q9Y4J3 PRELIMINARY; PRT; 8 AA.  
AC Q9Y4J3;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE RUNT/68NT/MTG8 PROTEIN (FRAGMENT).  
GN RUNT/68NT/MTG8.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE-BONE MARROW;  
RX MEDLINE=95002916; Pubmed=7919324;  
RA Tighe J.E., Calabi F.;  
RT "Alternative, out-of-frame runt/MTG8 transcripts are encoded in the  
RT derivative (8) chromosome in the t(8;21) of acute myeloid leuk.  
RT M2.;  
RL Blood 84:2115-2121(1994).  
DR EMBL; S74094; AAD14973.2;  
FT NON\_TER 1  
SQ SEQUENCE 8 AA; 929 MW; 30B764405B17244B CRC64;

Query Match 42.0%; Score 21; DB 4; Length 8;  
Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENMERN 6  
Db 1 ENLEIN 6

RESULT 2  
Q9Y4J4 PRELIMINARY; PRT; 8 AA.  
AC Q9Y4J4;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE RUNT/82NT/MTG8 PROTEIN (FRAGMENT).  
 GN RUNT/82NT/MTG8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95002916; PubMed=7919324;  
 RA Tighe J.E., Calabi F.;  
 RT "Alternative, out-of-frame runt/MTG8 transcripts are encoded by the  
 RT derivative (8) chromosome in the t(8;21) of acute myeloid leukemia  
 RT M2.";  
 RL Blood 84:2115-2121(1994).  
 DR EMBL: S74092; AAD14144.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 1067 MW; 20F414044B17244B CRC64;  
  
 Query Match 36.0%; Score 18; DB 4; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 ENME 4  
 Db 1 ENLE 4  
  
 RESULT 3  
 ID Q48686 PRELIMINARY; PRT; 9 AA.  
 AC Q48686;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE PROMOTER 23 DNA FRAGMENT (FRAGMENT).  
 OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Lactococcus;  
 OX NCBI\_TaxID=1359;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88105390; PubMed=2447829;  
 RA van der Vossen J.M., der Lelie D., Venema G.;  
 RT "Isolation and characterization of Streptococcus cremoris wg2-specific  
 RT promoters.";  
 RL Appl. Environ. Microbiol. 53:2452-2457(1987).  
 DR EMBL: M24763; AAA74720.1; -.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1080 MW; 5AF3A44AA4469443 CRC64;  
  
 Query Match 32.0%; Score 16; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 Qy 1 ENMERN 6  
 Db 2 KNMNDN 7  
  
 RESULT 4  
 ID Q9SB24 PRELIMINARY; PRT; 8 AA.  
 AC Q9SB24;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE PARAT PROTEIN (FRAGMENT).  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;

OC Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Niwa Y., Muranaka T., Baba A., Machida Y.;  
 RT "Organ-specific and auxin-inducible expression of two tobacco parA-  
 RT related genes in transgenic plants.";  
 RL DNA Res. 0:0-0(1994).  
 DR EMBL: D42119; BAA07700.1; -.  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 905 MW; FE32D2C44455BB16 CRC64;  
  
 Query Match 30.0%; Score 15; DB 10; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 3 MERN 6  
 Db 1 MESN 4  
  
 RESULT 5  
 ID Q85562 PRELIMINARY; PRT; 8 AA.  
 AC Q85562;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE TRUNCATED ENV PROTEIN (FRAGMENT).  
 OS Moloney murine leukemia virus.  
 OC Viruses; Retroviruses; Retroviridae; Mammalian type C retroviruses.  
 OX NCBI\_TaxID=11801;  
 RN [1]  
 RP SEQUENCE OF 4-8 FROM N.A.  
 RX MEDLINE=82196891; PubMed=6281735;  
 RA Donoghue D.J., Hunter T.;  
 RT "A generalized method of subcloning DNA fragments by restriction site  
 RT reconstruction: application to sequencing the amino-terminal coding  
 RT region of the transforming gene of Gazdar murine sarcoma virus.";  
 RL Nucleic Acids Res. 10:2549-2564(1982).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83164305; PubMed=6300424;  
 RA Donoghue D.J., Hunter T.;  
 RT "Recombinational junctions of variants of Moloney murine sarcoma  
 RT virus: generation and divergence of a mammalian transforming gene.";  
 RL J. Virol. 45:607-617(1983).  
 DR EMBL: K03105; AAA46490.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 732 MW; 98C2D5BEB44DC76D CRC64;  
  
 Query Match 30.0%; Score 15; DB 12; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 6 NC 7  
 Db 4 NC 5  
  
 RESULT 6  
 ID Q16386 PRELIMINARY; PRT; 9 AA.  
 AC Q16386;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE MEX40 PROTEIN (FRAGMENT).  
 GN MEX40.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95400293; PubMed=7670464;  
 RA Budarf M.L., Collins J., Gong W., Roe B., Wang Z., Bailey L.C.,  
 RA Sellinger B., Michaud D., Driscoll D.A., Emanuel B.S.;  
 RT "Cloning a balanced translocation associated with DiGeorge syndrome  
 RT and identification of a disrupted candidate gene.";  
 RL Nat. Genet. 10:269-278(1993).  
 DR EMBL; S79485; AADI4301.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1137 MW; 734911A69446837B CRC64;

Query Match 30.0%; Score 15; DB 4; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NMER 5  
 Db [1] 1  
 5 NMTR 8

RESULT 7  
 Q16276 PRELIMINARY; PRT; 9 AA.

AC Q16276;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE NEUROFIBROMATOSIS TYPE 2 (FRAGMENT).  
 GN NF2.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=95038750; PubMed=7951241;  
 RA Arai E., Ikeuchi T., Nakamura Y.;  
 RT "Characterization of the translocation breakpoint on chromosome  
 RT 22q12.2 in a patient with neurofibromatosis type 2 (NF2).";  
 RL Hum. Mol. Genet. 3:937-939(1994).  
 DR EMBL; S75841; AADI4190.2; -;  
 FT NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1044 MW; 1E4CBDD337232404 CRC64;

Query Match 30.0%; Score 15; DB 4; Length 9;  
 Best Local Similarity 42.9%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 MERNRA 9  
 Db [1] 1  
 1 MNKLLA 7

RESULT 8  
 Q67605 PRELIMINARY; PRT; 9 AA.

AC Q67605;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE A COMPONENT DNA (FRAGMENT).  
 GN ALL.

OS Squash leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10829;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=91082450; PubMed=1984669;  
 RA Lazarowitz S.G.;  
 RT "Molecular characterization of two bipartite geminiviruses causing  
 RT squash leaf curl disease: role of viral replication and movement  
 RT functions in determining host range.";  
 RL Virology 180:70-80(1991).  
 DR EMBL; M63155; AAA47823.1; -;  
 FT NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1118 MW; 2B30D5B457645417 CRC64;

Query Match 30.0%; Score 15; DB 12; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 MERN 6  
 Db [1] 1  
 1 MPRN 4

RESULT 9  
 Q67606 PRELIMINARY; PRT; 9 AA.

AC Q67606;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE A COMPONENT DNA (FRAGMENT).  
 GN ALL.

OS Squash leaf curl virus.  
 OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
 OX NCBI\_TaxID=10829;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=91082450; PubMed=1984669;  
 RA Lazarowitz S.G.;  
 RT "Molecular characterization of two bipartite geminiviruses causing  
 RT squash leaf curl disease: role of viral replication and movement  
 RT functions in determining host range.";  
 RL Virology 180:70-80(1991).  
 DR EMBL; M63157; AAA47822.1; -;  
 FT NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1118 MW; 2B30D5B457645417 CRC64;

Query Match 30.0%; Score 15; DB 12; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 MERN 6  
 Db [1] 1  
 1 MPRN 4

RESULT 10  
 Q9YVE3 PRELIMINARY; PRT; 7 AA.

AC Q9YVE3;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE PVI CORE PROTEIN (FRAGMENT).  
 GN PVI.

OS Human adenovirus type 7.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=10519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GOMEN;  
 RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;  
 RT "Molecular surveillance of strain variation in adenoviruses causing  
 RT acute respiratory disease, AV 4 and AV 7a.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF065065; AAD03662.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 28.0%; Score 14; DB 12; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 RNC 7  
 Db 4 RRC 6

RESULT 11  
 Q9YI0 Q9YI0 PRELIMINARY; PRT; 7 AA.  
 AC Q9YI0;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE PVI CORE PROTEIN (FRAGMENT).  
 GN PVI.  
 OS Human adenovirus type 7a.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=85755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KN T96-0620, S-1058, CL 68578;  
 RA Crawford-Miksz L.K., Nang R.N., Schnurr D.P.;  
 RT "Molecular surveillance of strain variation in adenoviruses causing acute respiratory disease, AV 4 and AV 7a."  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF065068; AAD03668.1; -  
 DR EMBL; AF065066; AAD03664.1; -  
 DR EMBL; AF065067; AAD03666.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 28.0%; Score 14; DB 12; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 RNC 7  
 Db 4 RRC 6

RESULT 12  
 Q9YI0 Q9YI0 PRELIMINARY; PRT; 7 AA.  
 AC Q9YI0;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE PVI CORE PROTEIN (FRAGMENT).  
 GN PVI.  
 OS Human adenovirus type 4.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=28280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2-G 95-873, RI-67, 55142;  
 RA Crawford-Miksz L.K., Nang R.N., Schnurr D.P.;  
 RT "Molecular surveillance of strain variation in adenoviruses causing acute respiratory disease, AV 4 and AV 7a."  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF065064; AAD03659.1; -  
 DR EMBL; AF065062; AAD03653.1; -  
 DR EMBL; AF065063; AAD03656.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 28.0%; Score 14; DB 12; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 RNC 7  
 Db 4 RRC 6

RESULT 13  
 Q31363 Q31363 PRELIMINARY; PRT; 9 AA.  
 AC Q31363;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSPC.  
 OS Borrelia garinii.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=29519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PBI;  
 RX MEDLINE=97426044; PubMed=9282748;  
 RA Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hovan D., Rosa P.;  
 RT "The Borrelia burgdorferi circular plasmid cp26: conservation of plasmid structure and targeted inactivation of the ospC gene."  
 RL Mol. Microbiol. 25:361-374(1997).  
 DR EMBL; U93699; AAC45533.1; -  
 DR NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1019 MW; 4864C1A731A44333 CRC64;

Query Match 28.0%; Score 14; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MERN 6  
 Db 1 MKN 4

RESULT 14  
 Q9R792 Q9R792 PRELIMINARY; PRT; 9 AA.  
 AC Q9R792;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE OUTER SURFACE PROTEIN C (FRAGMENT).  
 GN OSPC.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B31;  
 RA Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hovan D., Rosa P.;  
 RT "The Borrelia burgdorferi circular plasmid cp26: conservation of plasmid structure and targeted inactivation of the ospC gene."  
 RL Mol. Microbiol. 25:361-374(1997).  
 DR EMBL; U93693; AAC45521.1; -  
 DR NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1005 MW; 4864C5B731A44333 CRC64;

Query Match 28.0%; Score 14; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;

Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MERN 6  
|::|  
Db 1 MKKN 4

## RESULT 15

Q9R3T0 PRELIMINARY; PRT; 9 AA.  
AC Q9R3T0: 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DE 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE OUTER SURFACE PROTEIN C (FRAGMENT).  
GN OSCP.  
OS Borrelia afzelii.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=29518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PKO, IP21, AND J1;  
RA Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,  
Rosa P.;  
RT "he Borrelia burgdorferi circular plasmid cp26: conservation of  
plasmid structure and targeted inactivation of the ospC gene.";  
RL Mol. Microbiol. 25:361-374(1997).  
DR EMBL; U93698; AAC45331.1; -.  
DR EMBL; U93696; AAC45327.1; -.  
DR EMBL; U93697; AAC45329.1; -.  
FT NON\_TER 9  
SQ SEQUENCE 9 AA; 1005 MW; 4864C5B731A44333 CRC64;

Query Match 28.0%; Score 14; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. NO. 3.7e+05;

Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MERN 6  
|::|  
Db 1 MKKN 4

Search completed: January 17, 2001, 13:44:43  
Job time: 197 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:42:57 ; Search time 31.16 Seconds  
(without alignments)  
5.187 Million cell updates/sec

Title: US-08-765-837-8  
Perfect score: 50  
Sequence: 1 ENMERNCRA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 44977

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/PCrUS\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	23	46.0	5	1	US-07-671-757-18
2	22	44.0	9	1	US-08-167-336A-6
3	22	44.0	9	1	US-08-416-962-6
4	22	44.0	9	1	US-08-589-011-6
5	22	44.0	9	2	US-08-948-762-6
6	21	42.0	8	3	US-09-074-912-11
7	21	42.0	9	1	US-07-841-662-25
8	21	42.0	9	1	US-08-209-797-25
9	21	42.0	9	1	US-08-146-145-12
10	21	42.0	9	1	US-08-054-860-1
11	21	42.0	9	2	US-08-100-691-1
12	21	42.0	9	2	US-08-669-685-25
13	21	42.0	9	2	US-08-700-035A-14
14	21	42.0	9	3	US-09-103-486-25
15	21	42.0	9	3	US-08-481-985B-138
16	21	42.0	9	3	US-08-442-378-1
17	21	42.0	9	4	PCT-US93-01557-25
18	21	42.0	9	4	PCT-US95-04975-13
19	21	42.0	9	4	PCT-US95-16415-6
20	21	42.0	9	4	PCT-US96-13457-14
21	20	40.0	4	2	US-08-441-871-38
22	20	40.0	5	1	US-07-671-757-6
23	20	40.0	5	1	US-07-671-757-17
24	20	40.0	6	3	US-08-915-189-82
25	20	40.0	6	3	US-08-972-760-82
26	20	40.0	6	3	US-09-089-645A-82
27	20	40.0	7	1	US-07-789-184-195
28	20	40.0	7	1	US-08-475-263-195

29	20	40.0	7	1	US-08-485-886-195	Sequence 195, App
30	20	40.0	7	2	US-08-477-362-195	Sequence 195, App
31	20	40.0	7	2	US-08-477-134-195	Sequence 195, App
32	20	40.0	7	3	US-08-473-489A-195	Sequence 195, App
33	20	40.0	7	3	US-08-485-695-195	Sequence 195, App
34	20	40.0	9	1	US-08-173-510B-66	Sequence 66, Appl
35	20	40.0	9	1	US-08-615-181-47	Sequence 47, Appl
36	20	40.0	9	1	US-08-458-218-64	Sequence 64, Appl
37	20	40.0	9	2	US-08-450-497-66	Sequence 25, Appl
38	19	38.0	5	1	US-07-671-757-25	Sequence 26, Appl
39	19	38.0	6	1	US-07-671-757-26	Sequence 31, Appl
40	19	38.0	6	1	US-08-049-794-31	Sequence 31, Appl
41	19	38.0	6	1	US-08-496-847-31	Sequence 31, Appl
42	19	38.0	6	2	US-08-742-774-31	Sequence 31, Appl
43	19	38.0	6	2	US-08-675-354-31	Sequence 31, Appl
44	19	38.0	6	2	US-08-965-918-31	Sequence 31, Appl
45	19	38.0	6	2	US-09-138-433-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1  
US-07-671-757-18  
; Sequence 18, Application US/07671757  
; Patent No. 5484770  
; GENERAL INFORMATION:  
; APPLICANT: LAERUM, Ole D  
; TITLE OF INVENTION: PEPTIDE COMPOUNDS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bacon & Thomas  
; STREET: 625 Slaters Lane - Fourth Floor  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States of America  
; ZIP: 22314  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/671,757  
; FILING DATE: 19910408  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8821785.6  
; FILING DATE: 16-SEP-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FICHTER, Richard E  
; REGISTRATION NUMBER: 26,382  
; REFERENCE/DOCKET NUMBER: REF/Laerum/757  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 683-0500  
; TELEFAX: (703) 683-1080  
; TELEX: 89-9124 BATO ACTN  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1  
; OTHER INFORMATION: /label= pyro  
; OTHER INFORMATION: /note= "pyroglutamine"  
US-07-671-757-18

Query Match 46.0%; Score 23; DB 1; Length 5;  
Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERNCR 8  
Db 1 EQNCK 5

RESULT 2  
US-08-167-336A-6  
; Sequence 6, Application US/08167336A  
; Patent No. 5531990  
; GENERAL INFORMATION:  
; APPLICANT: THANAVALA, YASMIN  
; APPLICANT: THAKUR, ARVIND  
; APPLICANT: ROITT, IVAN  
; APPLICANT: PRIDE, MICHAEL  
; TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY  
; TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS  
; TITLE OF INVENTION: B SURFACE ANTIGEN  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DUNN & ASSOCIATES, P.C.  
; STREET: P.O. BOX 96  
; CITY: NEWFANE  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 14108  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB  
; COMPUTER: VICTOR 300 SX/25  
; OPERATING SYSTEM: MS-DOS VERSION 5.0  
; SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/167,336A  
; FILING DATE: 15-DEC-1993  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DUNN, MICHAEL L.  
; REGISTRATION NUMBER: 25,330  
; REFERENCE/DOCKET NUMBER: RPP:138 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 433-1661  
; TELEFAX: (716) 433-1665  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: AMINO ACID  
; STRANDEDNESS: UNKNOWN  
; TOPOLOGY: UNKNOWN  
; MOLECULE TYPE: PEPTIDE  
; HYPOTHETICAL: NO  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:

MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-167-336A-6

Query Match 44.0%; Score 22; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERNC 7  
Db 6 DRNC 9

RESULT 3  
US-08-416-962-6  
; Sequence 6, Application US/03416962  
; Patent No. 5668253  
; GENERAL INFORMATION:  
; APPLICANT: THANAVALA, YASMIN  
; APPLICANT: THAKUR, ARVIND  
; APPLICANT: ROITT, IVAN  
; APPLICANT: PRIDE, MICHAEL  
; TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY  
; TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS  
; TITLE OF INVENTION: B SURFACE ANTIGEN  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DUNN & ASSOCIATES, P.C.  
; STREET: P.O. BOX 96  
; CITY: NEWFANE  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 14108  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB  
; COMPUTER: VICTOR 300 SX/25  
; OPERATING SYSTEM: MS-DOS VERSION 5.0  
; SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/416,962  
; FILING DATE: 05-APR-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/167,336  
; FILING DATE: 15-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DUNN, MICHAEL L.  
; REGISTRATION NUMBER: 25,330  
; REFERENCE/DOCKET NUMBER: RPP:138 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 433-1661  
; TELEFAX: (716) 433-1665  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9

TYPE: AMINO ACID  
STRANDEDNESS: UNKNOWN  
TOPOLOGY: UNKNOWN  
MOLECULE TYPE: PEPTIDE  
HYPOTHETICAL: NO  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:

STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:

CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:

FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:

TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO:  
US-08-416-962-6

Query Match 44.0%; Score 22; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERNC 7  
Db 6 DRNC 9

RESULT 4

US-08-589-011-6  
Sequence 6, Application US/08589011  
Patent No. 5744135

GENERAL INFORMATION:

APPLICANT: THANAVALA, YASMIN  
APPLICANT: THAKUR, ARVIND  
APPLICANT: ROITT, IVAN  
APPLICANT: PRIDE, MICHAEL

TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY  
TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS  
TITLE OF INVENTION: B SURFACE ANTIGEN  
NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: DUNN & ASSOCIATES, P.C.  
STREET: P.O. BOX 96

CITY: NEWFANE

STATE: NEW YORK

COUNTRY: USA

ZIP: 14108

COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB  
COMPUTER: VICTOR 300 SX/25  
OPERATING SYSTEM: MS-DOS VERSION 5.0  
SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/589,011  
FILING DATE: 19-JAN-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/167,336  
FILING DATE: 15-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: DUNN, MICHAEL L.  
REGISTRATION NUMBER: 25,330  
REFERENCE/DOCKET NUMBER: RPP:138 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 433-1661  
TELEFAX: (716) 433-1665  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: AMINO ACID  
STRANDEDNESS: UNKNOWN  
TOPOLOGY: UNKNOWN  
MOLECULE TYPE: PEPTIDE  
HYPOTHETICAL: NO  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

US-08-589-011-6

Query Match 44.0%; Score 22; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ERNC 7  
Db 6 DRNC 9

```
RESULT 5
US-08-948-762-6
; Sequence 6, Application US/08948762
; Patent No. 5856087
; GENERAL INFORMATION:
; APPLICANT: THANAVALA, YASMIN
; APPLICANT: THAKUR, ARVIND
; APPLICANT: ROITT, IVAN
; APPLICANT: PRIDE, MICHAEL
; TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODY
; TITLE OF INVENTION: HAVING CORRESPONDENCE WITH HUMAN HEPATITIS
; TITLE OF INVENTION: B SURFACE ANTIGEN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DUNN & ASSOCIATES, P.C.
; STREET: P.O. BOX 96
; CITY: NEWFANE
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 14108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: VICTOR 300 SX/25
; OPERATING SYSTEM: MS-DOS VERSION 5.0
; SOFTWARE: WORDSTAR PROFESSIONAL RELEASE 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,762
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; FILING APPLICATION NUMBER: 08/589,011
; FILING DATE: 19-JAN-1996
; APPLICATION NUMBER: 08/167,336
; FILING DATE: 15-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DUNN, MICHAEL L.
; REGISTRATION NUMBER: 25,330
; REFERENCE/DOCKET NUMBER: RPP-138 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 433-1661
; TELEFAX: (716) 433-1665
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9
; TYPE: AMINO ACID
; STRANDEDNESS: UNKNOWN
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE: PEPTIDE
; HYPOTHETICAL: NO
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:

; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-08-948-762-6

Query Match 44.0%; Score 22; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Caps 0;

QY 4 ERNC 7
Db 6 DRNC 9

RESULT 6
US-09-074-912-11
; Sequence 11, Application US/09074912
; Patent No. 6057491
; GENERAL INFORMATION:
; APPLICANT: CIGAN, AMY L
; APPLICANT: CZAPLA, THOMAS H
; APPLICANT: FALLIS, LYNN E
; APPLICANT: MEYER, TERRY E
; APPLICANT: MUNDELL, SCOTT A
; APPLICANT: SABUS, BRIAN
; APPLICANT: SCHUBERT, KAREL
; TITLE OF INVENTION: PROTEINS HAVING INSECTICIDAL ACTIVITIES
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. MURRAY SPRUILL (ALSTON & BIRD, LLP)
; STREET: 3605 GLENWOOD AVE.
; CITY: RALEIGH
; STATE: NC
; COUNTRY: USA
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,912
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: SPRUILL, W. MURRAY
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Pentaclethra macroloba
```



US-09-074-912-11

Query Match 42.0%; Score 21; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENNE 4  
||||  
DB 1 ENNE 4

## RESULT 7

US-07-841-662-25  
; Sequence 25, Application US/07841662  
; Patent No. 5314813  
; GENERAL INFORMATION:  
; APPLICANT: Peterson, Per A  
; APPLICANT: Jackson, Michael  
; APPLICANT: Lengade-Demoyen, Pierre  
; TITLE OF INVENTION: IN VITRO ACTIVATION OF CYTOTOXIC T CELLS  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute  
; STREET: 10666 No. 5314813th Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/841,662  
; FILING DATE: 19920219  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Logan, April  
; REGISTRATION NUMBER: 33,950  
; REFERENCE/DOCKET NUMBER: SPF0001P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 554-2937  
; TELEFAX: (619) 554-6312  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-07-841-662-25

Query Match 42.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENNE 4  
||||  
DB 4 ENNE 7

## RESULT 8

US-08-209-797-25  
; Sequence 25, Application US/08209797

Patent No. 5529921  
; GENERAL INFORMATION:  
; APPLICANT: Peterson, Per A  
; APPLICANT: Jackson, Michael  
; APPLICANT: Lengade-Demoyen, Pierre  
; TITLE OF INVENTION: IN VITRO ACTIVATION OF CYTOTOXIC T CELLS  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute  
; STREET: 10666 No. 5529921th Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/209,797  
; FILING DATE: 10-MAR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/841,662  
; FILING DATE: 19-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Logan, April  
; REGISTRATION NUMBER: 33,950  
; REFERENCE/DOCKET NUMBER: SPF0001P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 554-2937  
; TELEFAX: (619) 554-6312  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-209-797-25

Query Match 42.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENNE 4  
||||  
DB 4 ENNE 7

## RESULT 9

US-08-146-145-12  
; Sequence 12, Application US/08146145  
; Patent No. 5747269  
; GENERAL INFORMATION:  
; APPLICANT: Rammensee, Hans-Georg  
; APPLICANT: Falk, Kirsten  
; APPLICANT: R tzsckke, Olaf  
; APPLICANT: Stevanovic, Stefan  
; APPLICANT: Jung, G nther  
; TITLE OF INVENTION: DETERMINATION OF PEPTIDE MOTIFS ON MHC  
; TITLE OF INVENTION: MOLECULES  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.

; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/146,145  
; FILING DATE: 17-NOV-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitts, Monica C.  
; REGISTRATION NUMBER: 36,105  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-146-145-12

Query Match 42.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENME 4  
; ; ; ; ;  
Db 4 ENME 7

RESULT 10  
US-08-054-860-1  
; Sequence 1, Application US/08054860  
; Patent No. 5807559  
; GENERAL INFORMATION:  
; APPLICANT: Jondal, Mikael  
; TITLE OF INVENTION: New Active Compounds  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/054,860  
; FILING DATE: 19930427  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9201338-2  
; FILING DATE: 28-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9202553-5  
; FILING DATE: 07-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9203897-5  
; FILING DATE: 23-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: SE 9301141-9  
; FILING DATE: 06-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Starnier, Richard J

; REGISTRATION NUMBER: P-35, 372  
; REFERENCE/DOCKET NUMBER: 1103326  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-054-860-1

Query Match 42.0%; Score 21; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENME 4  
; ; ; ; ;  
Db 4 ENME 7

RESULT 11  
US-08-100-691-1  
; Sequence 1, Application US/08100691  
; Patent No. 5820872  
; GENERAL INFORMATION:  
; APPLICANT: Edelson, Richard L.  
; APPLICANT: Gasparro, Francis P.  
; TITLE OF INVENTION: Cellular Vaccine and Methods of Use for  
; TITLE OF INVENTION: the Treatment of Solid Tumor Malignancies  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Ave.  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/100,691  
; FILING DATE: 19930730  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977672  
; FILING DATE: 18-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: Y006077005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/720-3500  
; TELEFAX: 617/720-2441  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-08-100-691-1

Query Match 42.0%; Score 21; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENME 4  
; ; ; ; ;

Db

4 ENME 7

## RESULT 12

US-08-669-685-25  
; Sequence 25, Application US/08669685  
; Patent No. 5827737  
; GENERAL INFORMATION:  
; APPLICANT: Peterson, Per A  
; APPLICANT: Jackson, Michael  
; APPLICANT: Lengade-Demoyen, Pierre  
; TITLE OF INVENTION: IN VITRO ACTIVATION OF CYTOTOXIC T CELLS  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute  
; STREET: 10666 No. 5827737th Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/669,685  
; FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/209,797

FILING DATE: 10-MAR-1994

APPLICATION NUMBER: US 07/841,662

FILING DATE: 19-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Logan, April

REGISTRATION NUMBER: 33,950

REFERENCE/DOCKET NUMBER: SPFO001P

TELEPHONE: (619) 554-2937

TELEFAX: (619) 554-6312

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE:

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

US-08-669-685-25

Query Match 42.0%; Score 21; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 ENME 4

Db

4 ENME 7

## RESULT 13

US-08-700-035A-14  
; Sequence 14, Application US/08700035A  
; Patent No. 5831068  
; GENERAL INFORMATION:  
; APPLICANT: Nair, et al., Smita K.  
; TITLE OF INVENTION: A METHOD TO INCREASE THE DENSITY OF  
; ANTIGEN ON ANTIGEN PRESENTING CELLS  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/700,035A

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/517,373

FILING DATE: 21-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 06765/009001

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-700-035A-14

Query Match 42.0%; Score 21; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 ENME 4

Db

4 ENME 7

## RESULT 14

US-09-103-486-25  
; Sequence 25, Application US/09103486  
; Patent No. 6001365  
; GENERAL INFORMATION:  
; APPLICANT: Peterson, Per A  
; APPLICANT: Jackson, Michael  
; APPLICANT: Lengade-Demoyen, Pierre  
; TITLE OF INVENTION: IN VITRO ACTIVATION OF CYTOTOXIC T CELLS  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute  
; STREET: 10666 No. 6001365th Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/103,486

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/669,685

;  
; FILING DATE: US 07/841,662  
; FILING DATE: 19-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Logan, April  
; REGISTRATION NUMBER: 33,950  
; REFERENCE/DOCKET NUMBER: SPF0001P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 554-2937  
; TELEFAX: (619) 554-6312  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-09-103-486-25

Query Match 42.0%; Score 21; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ENME 4  
Db 4 ENME 7

RESULT 15  
US-08-481-985B-138  
; Sequence 138, Application US/08481985B  
; Patent No. 6011146  
; GENERAL INFORMATION:  
; APPLICANT: Mottez, Estelle  
; APPLICANT: Abastado, Jean-Pierre  
; APPLICANT: Kourilsky, Philippe  
; TITLE OF INVENTION: Altered Major Histocompatibility Complex  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,985B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/801,818  
; FILING DATE: 05-DEC-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/792,473  
; FILING DATE: 15-NOV-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0106-04000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000

;  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 138:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-481-985B-138

Query Match 42.0%; Score 21; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ENME 4  
Db 4 ENME 7

Search completed: January 17, 2001, 13:42:58  
Job time: 157 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:42:23 ; Search time 43.61 Seconds

(without alignments)  
7.057 Million cell updates/sec

Title: US-08-765-837-9

Perfect score: 52

Sequence: 1 DPTFKENYR 9

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 61695

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*  
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	52	100.0	9	R90953	Factor VIII antigen
2	29	55.8	9	R6651	Ras(32-40) oncopro
3	29	55.8	9	W3467	Effector loop doma
4	29	55.8	9	W7989	Ras and Rap family
5	29	55.8	9	W5746	Ras-binding fragme
6	29	55.8	9	W4822	Raf-binding fragme
7	24	46.2	5	R37031	Pentapeptide compo
8	24	46.2	9	W71343	Glycolipid sugar c
9	23	44.2	6	R30572	Anaphylatoxin anta
10	23	44.2	7	R55100	Prostate-specific
11	23	44.2	7	W02239	Prostate-specific
12	23	44.2	9	Y55970	Human SULU1-deri

13	22	42.3	6	13	R24699	Immunomodulatory p
14	22	42.3	8	13	R27999	Tryptic peptide tr
15	22	42.3	9	15	Y37952	Human cdk2 onco
16	22	42.3	9	15	R48041	Neuropeptide Y ant
17	22	42.3	9	15	R48045	Neuropeptide Y ant
18	22	42.3	9	15	R48047	Neuropeptide Y ant
19	22	42.3	9	15	R59108	Peptide fragment (
20	22	42.3	9	19	W70054	HER-2/neu deriv
21	22	42.3	9	20	Y45520	Immunogenic peptid
22	21	40.4	6	19	W58999	Human HLA-B27 pept
23	21	40.4	7	19	W85927	Rodent IL-1 epsilo
24	21	40.4	8	14	R43412	La/SSB epitope 182
25	21	40.4	8	20	Y33320	CI12-ANG construct
26	21	40.4	9	6	P50204	Sequence of oligop
27	21	40.4	9	14	R41804	Thrombomodulin pep
28	21	40.4	9	16	R71600	Solanum tuberosum
29	21	40.4	9	20	W97577	Peptide sequence p
30	20	38.5	5	14	R37032	Pentapeptide compo
31	20	38.5	6	11	R04175	Monomer of repetat
32	20	38.5	6	16	R72787	Yeast ribonucleoti
33	20	38.5	6	20	W92268	Human anti-angiot
34	20	38.5	6	21	Y83893	Yeast ribonucleoti
35	20	38.5	7	11	R07931	Tryptic fragment T
36	20	38.5	7	11	R07651	Ribonuclease reduc
37	20	38.5	7	16	R72786	Yeast ribonucleoti
38	20	38.5	7	17	R94379	Patched protein co
39	20	38.5	7	18	W12028	Herpes virus DNA p
40	20	38.5	7	19	W53116	Peptide 3 of unspe
41	20	38.5	7	21	Y83839	Ribonucleotide red
42	20	38.5	7	21	Y83892	Yeast ribonucleoti
43	20	38.5	7	21	Y62946	PA-cadherin cell a
44	20	38.5	8	9	R83008	Blocking peptide u
45	20	38.5	8	21	Y73038	Hepatitis B virus

## ALIGNMENTS

RESULT 1  
R90953  
ID R90953 standard; peptide; 9 AA.  
XX  
AC R90953;  
XX  
DT 09-SEP-1996 (first entry)  
XX  
DE Factor VIII antigenic peptide corresp. to residues Aspl909-Arg1917.  
KW Factor VIII: modification; inhibitor activity; binding; antibody;  
KW von Willebrand factor; immune disorder.  
XX  
OS Synthetic.  
XX  
PN WO9602572-A2.  
XX  
PD 01-FEB-1996.  
XX  
PF 14-JUL-1995; 95WO-BE000068.  
XX  
PR 14-JUL-1994; 94BE-0000666.  
XX  
PA (CROI-) CROIX ROUGE BELGIQUE.  
PI Di Giambattista M, Laub R;  
DR WPI; 1996-105861/11.  
PT Factor VIII antigenic polypeptide fragments and epitope(s) - also  
PT inhibitors of factor VIII and anti-inhibitors, useful for e.g.  
PT preventing and treating immune disorders involving inhibition of  
XX factor VIII binding  
PS Claim 8; Page 13; 45pp; French.

XX Peptides R90945-64 are derived from the factor VIII protein, esp. from a  
 CC modified Factor VIII in which residues Ala322-Ser750, Leu1655-Arg1689,  
 CC Lys1694-Pro1782 and Asp2170-Tyr2332 are deleted. The modified Factor VIII  
 CC and derived peptides can be used to prevent the activity of inhibitors of  
 CC factor VIII binding to von Willebrand factor, esp. antibodies, thus  
 CC preventing or treating immune disorders.  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 17; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPTFKENYR 9  
 DB 1 dptfkenyr 9

RESULT 2  
 R66651  
 ID R66651 standard; Peptide; 9 AA.

XX AC R66651;

XX DT 03-AUG-1995 (first entry)

XX DE Ras(32-40) oncoprotein fragment.

XX KW Oncoprotein; oncogene protein; Raf; Ras; zinc-finger domain;  
 KW cancer; therapy; autoimmune disease; anti-proliferative; antitumor;  
 KW anti-oncogene; tumor suppressor.

XX OS Synthetic.

XX PN W09429727-A.

XX PD 22-DEC-1994.

XX PF 10-JUN-1994; 94WO-US06654.

XX PR 11-JUN-1993; 93US-0077256.

XX PA (GEO ) GEN HOSPITAL CORP.

XX PA (INDV ) UNIV INDIANA FOUND.

XX PI Avruch J, Marshall MS, Zhang X;

XX DR WPI; 1995-036668/05.

XX PT New peptide(s) which inhibit Ras interaction with Raf - useful  
 PT for treating cancer and autoimmune disease

XX PS Claim 12; Page 66; 94pp; English.

XX CC Compounds with putative anti-oncogene or tumor suppressor activity  
 CC are assayed by their ability to bind a Ras-binding fragment of Raf  
 CC oncoprotein, such as the peptides given in R66644-49 and R66655-56,  
 CC or to bind a Raf-binding fragment of Ras, such as the peptides given  
 CC in R66650-52.

XX SQ Sequence 9 AA;

Query Match 55.8%; Score 29; DB 16; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFKENY 8  
 DB 2 dptfkeny 9

RESULT 3

W34647

ID W34647 standard; peptide; 9 AA.

XX AC W34647;

XX DT 01-APR-1998 (first entry)

XX DE Effector loop domain (amino acids 32-40) of Ras.

XX KW Raf protein; effector loop domain; Ras; signal transduction;  
 KW Ras-Raf mediated signal transduction pathway; cell proliferation;  
 KW binding; inhibition; Raf activation inhibiting activity;  
 KW cell proliferation inhibition; immune cell receptor-mediated activation.

XX OS Unidentified.

XX PN W09734146-A1.

XX PD 18-SEP-1997.

XX PF 12-MAR-1997; 97WO-US03881.

XX PR 11-MAR-1997; 97US-0814836.

XX PR 12-MAR-1996; 96US-0013274.

XX PA (GEO ) GEN HOSPITAL CORP.

XX PA (INDV ) UNIV INDIANA FOUND.

XX PI Avruch J, Luo Z, Marshall MS;

XX DR WPI; 1997-470979/43.

XX PT Screening compounds which inhibit direct protein interaction -  
 PT useful to identify inhibitors of binding of Ras to Raf, Raf  
 PT activation and cell proliferation

XX PS Claim 9; Page 3; 50pp; English.

XX CC The present sequence represents amino acids 32-40 of Ras, which comprise  
 CC the effector loop domain. The effector loop domain binds to amino  
 CC acids 50-150 (W34649) of the Raf protein. This interaction participates  
 CC in the transduction of an intracellular signal via the Ras-Raf mediated  
 CC signal transduction pathway which culminates in cell proliferation. A  
 CC novel method for reducing cell proliferation in a mammal comprises  
 CC administering to or contacting the cells with a compound which inhibits  
 CC direct binding of the effector loop domain of Ras with an amino  
 CC terminal Ras binding domain of Raf. The candidate compound may contain  
 CC the present sequence. The methods are used to screen compounds for Raf  
 CC activation inhibiting activity, cell proliferation inhibition and  
 CC inhibition of direct binding of Ras to Raf. Blocking association of Ras  
 CC with Raf interferes with receptor-mediated activation of immune cells.  
 CC The method may also be useful in down-regulating the immune response in  
 CC patients with autoimmune diseases such as systemic lupus erythematosus,  
 CC type 1 diabetes, and rheumatoid arthritis.

XX SQ Sequence 9 AA;

Query Match 55.8%; Score 29; DB 18; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFKENY 8  
 DB 2 dptfkeny 9

RESULT 4  
 W79899  
 ID W79899 standard; Peptide; 9 AA.  
 XX

AC W79899;  
 XX  
 DT 18-JAN-1999 (first entry)  
 XX  
 DE Ras and Rap family effector domain.  
 XX  
 KW NOEY2; tumour suppressor; human; breast cancer; ovary cancer;  
 KW diagnosis; therapy; transgenic animal; vaccine; Ras; Rap;  
 KW effector domain.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09842830-A2.  
 XX  
 PD 01-OCT-1998.  
 XX  
 PF 20-MAR-1998; 98WO-US05723.  
 XX  
 PR 13-JAN-1998; 98US-0071263.  
 PR 21-MAR-1997; 97US-0041580.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Bast RC, Xu F, Yu Y;  
 XX  
 DR WPI; 1998-532005/45.  
 XX  
 XX New nucleic acid encoding NOEY2 tumour suppressor from ovarian  
 PT epithelium - useful for, e.g. treatment, diagnosis and prognosis of  
 PT cancer, particularly cancer of ovary and breast  
 XX  
 PS Example 1; Page 102; 182pp; English.  
 XX  
 CC This is the amino acid sequence of the effector domain of Ras and  
 CC Rap family members. A putative effector domain of NOEY2 (see  
 CC W79899) differs in 3 residues from this sequence. NOEY2 is a novel  
 CC human tumour suppressor protein that is absent or down-regulated in  
 CC breast and ovarian cancers. NOEY2 polynucleotides (see V60577-78)  
 CC and polypeptides of the invention are used in the treatment,  
 CC diagnosis and prognosis of cancer, particularly of the ovary and  
 CC breast.  
 XX  
 SQ Sequence 9 AA;

Query Match 55.88; Score 29; DB 19; Length 9;  
 Best Local Similarity 50.08; Pred. No. 2.le+05;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFKENY 8  
 III : : I  
 Db 2 dptiedsy 9

RESULT 5  
 W57464  
 ID W57464 standard; peptide; 9 AA.  
 XX  
 AC W57464;  
 XX  
 DT 13-AUG-1998 (first entry)  
 XX  
 DE Ras-binding fragment of Raf protein Raf(32-40) (residues 32-40).  
 XX  
 KW Raf; Ras; Ras-binding fragment; inhibition; tumour; autoimmune disease;  
 KW dysregulated signal transduction; cell proliferation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5767075-A.  
 XX  
 PD 16-JUN-1998.  
 XX

PF 02-JUN-1995; 95US-0460533.  
 XX  
 PR 10-JUN-1994; 94US-0259672.  
 PR 11-JUN-1993; 93US-0077256.  
 PR 02-JUN-1995; 95US-0460533.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (INDV ) UNIV INDIANA FOUND.  
 XX  
 PI Avruch J, Marshall MS, Zhang X;  
 XX  
 DR WPI; 1998-361751/31.  
 XX  
 PT Inhibition of Ras-Raf interaction in vivo - by administration of Raf  
 PT peptide fragments  
 XX  
 PS Disclosure; Column 2; 35pp; English.  
 XX  
 CC This represents a Ras-binding peptide fragment of the Raf protein. The  
 CC Ras-binding peptides especially those shown in W57456 to W57460 can be  
 CC used in methods of inhibiting cell proliferation on an animal which  
 CC comprises inhibiting direct interaction of Ras with Raf by administering  
 CC the peptides. The Raf peptide can be used in methods of evaluating an  
 CC anti-proliferative compound by contacting the compound with a Raf and  
 CC determining its ability to bind to Ras. The peptides are useful for  
 CC treating diseases characterised by dysregulated signal transduction or  
 CC aberrant cell proliferation, e.g. tumours and autoimmune diseases.  
 XX  
 SQ Sequence 9 AA;

Query Match 55.88; Score 29; DB 19; Length 9;  
 Best Local Similarity 50.08; Pred. No. 2.le+05;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFKENY 8  
 III : : I  
 Db 2 dptiedsy 9

RESULT 6  
 W48229  
 ID W48229 standard; peptide; 9 AA.  
 XX  
 AC W48229;  
 XX  
 DT 18-JUN-1998 (first entry)  
 XX  
 DE Raf-binding fragment of Ras SEQ ID NO:4.  
 XX  
 KW Ras; Raf; binding peptide; oncogene; rat sarcoma virus; human;  
 KW inhibit; cell proliferation; tumour; autoimmune disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5736337-A.  
 XX  
 PD 07-APR-1998.  
 XX  
 PF 10-JUN-1994; 94US-0259672.  
 XX  
 PR 10-JUN-1994; 94US-0259672.  
 PR 11-JUN-1993; 93US-0077256.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (INDV ) UNIV INDIANA FOUND.  
 XX  
 PI Avruch J, Marshall MS, Zhang X;  
 XX  
 DR WPI; 1998-239205/21.  
 XX  
 PT Screening assays for anti-proliferative agents - is based on effect  
 PT of Raf protein binding

XX Claim 10; Column 35-36; 34pp; English.

XX The present sequence represents a Raf-binding peptide fragment of Ras, from the present invention. The present invention describes: (1) a method of evaluating an anti-proliferative compound comprising contacting the compound with an amino-terminal, non-catalytic Ras-binding fragment of Raf and determining the ability of the compound to bind to the fragment, where the binding ability of the compound is an indication that the compound inhibits cell proliferation; and (2) a method of evaluating an anti-proliferative compound comprising contacting the compound with a Raf-binding fragment of Ras comprising an intact effector loop and determining the ability of the compound to bind to the fragment, where the binding ability of the compound is an indication that the compound inhibits cell proliferation. Antiproliferative compounds identified as above can be used to treat tumours and autoimmune diseases.

XX Sequence 9 AA;

Query Match 55.8%; Score 29; DB 19; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.1e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFKENY 8  
   | | | : : : |  
Db 2 dptiedsy 9

RESULT 7  
ID R37031 standard; peptide; 5 AA.  
AC R37031;  
XX 28-JUL-1993 (first entry)  
XX Pentapeptide component #1 of N-linked glycoconjugate.  
DE beta-anomer; beta-glycosylamine linked glycoconjugate.  
KW Synthetic.  
XX EP538230-A.  
XX 21-APR-1993.  
XX 14-OCT-1992; 92EP-0870165.  
XX 15-OCT-1991; 91US-0776911.  
PR 11-AUG-1992; 92US-0926786.  
XX (MONS ) MONSANTO CO.  
XX Dwek RA, Manger ID, Rademacher TW, Wong SYC, Wong S;  
XX WPI; 1993-128045/16.  
XX N-linked peptide glyco-conjugate(s) prepn. - by reacting oligosaccharide(s) with ammonium bi:carbonate to maintain beta-anomeric configuration, and avoid sepn. of anomers  
PS Claim 3; Page 32; 50pp; English.

XX A peptide of 5-25 amino acid residues (pref. one of the pentapeptides R37031-3 or an atriopentin e.g. atriopentin A or D, see R37337-8) having an activated carboxyl group capable of forming a beta-glycosylamine linked glycoconjugate is reacted with an unprotected beta-glycosylamine deriv. of an oligosaccharide. The deriv. is prepared by reacting an unprotected oligosaccharide having up to 9 saccharide units with saturated ammonium bicarbonate at pH 8.0-8.5. The N-linked glycoconjugate is produced under conditions to directly maintain the beta-anomeric configuration.

XX Sequence 5 AA;

Query Match 46.2%; Score 24; DB 14; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.1e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPTFF 4  
   | | | |  
Db 2 dptf 5

RESULT 8  
ID W71343 standard; peptide; 9 AA.  
XX AC W71343;  
XX 26-NOV-1998 (first entry)  
XX Glycolipid sugar chain replica peptide #LC07.  
XX Glycolipid sugar chain replica peptide; antibody; control;  
KW glycosidase activity; treatment; liver disease.  
XX QS Synthetic.  
XX JPI0237098-A.  
XX 08-SEP-1998.  
XX 26-FEB-1997; 97JP-0042317.  
XX 26-FEB-1997; 97JP-0042317.  
XX (IMMO ) IMMUNO JAPAN INC.  
XX WPI; 1998-537487/46.  
XX Glyco-lipid sugar chain replica peptide - reacts specifically with antibody against glyco-lipid sugar chain and controls glycosidase activity, useful for improving treatment of liver diseases  
PS Claim 12; Page 4; 11pp; Japanese.

XX W71337-49 represent glycolipid sugar chain replica peptides that react specifically with an antibody against glycolipid sugar chain and controls glycosidase activity. The drug composition containing the peptides may be used to improve treatment for liver diseases.

XX Sequence 9 AA;

Query Match 46.2%; Score 24; DB 19; Length 9;  
Best Local Similarity 57.1%; Pred. No. 2.1e+05;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PTFKENY 8  
   | | | : |  
Db 3 ptfndvy 9

RESULT 9  
ID R30572 standard; peptide; 6 AA.  
XX AC R30572;  
XX 29-JAN-1993 (first entry)  
XX Anaphylatoxin antagonist or agonist peptide.  
XX



KW Anaphylatoxin; receptor; ligand; antiinflammatory; hexapeptide;  
 KW heptapeptide; immunodeficiency; allergy; autoimmune; ARDS; cancer;  
 KW infection; endotoxin; asthma; gout; psoriasis; cirrhosis;  
 KW inflammatory; bowel; disease; hepatitis; burns; myocardial;  
 KW infarction; transplant rejection; ischaemic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-Me-Phe"

FT Modified-site 4 /label= OTHER

FT Modified-site 6 /note= "(2R)-2-amino-3-cyclohexyl-propanoyl"

FT Modified-site 6 /note= "D-Arg"

FT Modified-site 6 /note= "D-Arg"

XX W09211858-A.

PN W09211858-A.

XX 23-JUL-1992.

PD 23-JUL-1992.

XX 10-DEC-1991; 91WO-US09319.

PF 10-DEC-1991; 91WO-US09319.

XX 27-DEC-1990; 90US-0634641.

PR 27-DEC-1990; 90US-0634641.

XX (ABBO ) ABBOTT LABORATORIES.

PA (ABBO ) ABBOTT LABORATORIES.

XX Kawai M, Lully JR, Or YS, Wagner R, Wiedeman PE;

PI Kawai M, Lully JR, Or YS, Wagner R, Wiedeman PE;

XX WPI; 1992-268383/32.

DR WPI; 1992-268383/32.

XX New hexa- and hepta-peptide(s) are anaphylatoxin antagonists and

PT agonists - for treating inflammatory and immunodeficiency

PT agonists - for treating inflammatory and immunodeficiency

PT agonists - for treating inflammatory and immunodeficiency

XX diseases, cancers and severe infections

PS Claim 11; Page 154; 160pp; English.

XX The peptide is a specifically claimed example of a group of highly

CC generic hexa- and heptapeptides which are (a) anaphylatoxin

CC antagonists useful for treating asthma, other allergies,

CC inflammations, autoimmune diseases, serum sickness, gout, bullous

CC skin diseases, psoriasis, ARDS, endotoxin shock, hepatic cirrhosis,

CC pancreatitis, inflammatory bowel disease, burns, sepsis, myocardial

CC infarction, chronic hepatitis, transplant rejection, or ischaemic

CC heart or brain damage; or (b) anaphylatoxin agonists useful for

CC stimulating inflammatory and immune responses, e.g. in the treatment

CC of cancer, immunodeficiency diseases and severe infections.

XX Sequence 6 AA;

Query Match

Best Local Similarity 44.2%; Score 23; DB 13; Length 6;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PKNYR 9

Db 1 Fkxfr 6

RESULT 10

R55100

ID R55100 standard; Peptide; 7 AA.

XX R55100;

AC R55100;

XX 11-JAN-1995 (first entry)

DT 11-JAN-1995 (first entry)

XX Prostate-specific membrane antigen peptide fragment.

DE Prostate-specific membrane antigen peptide fragment.

XX Prostate-specific membrane antigen; PSM; prostate cancer;

KW transmembrane glycoprotein; imaging; targeting; tumour detection;

KW antibody detection; sequencing.

XX Homo sapiens.

OS Homo sapiens.

XX W09409820-A.

PN W09409820-A.

XX 11-MAY-1994.

PD 11-MAY-1994.

XX 05-NOV-1993; 93WO-US10624.

PF 05-NOV-1993; 93WO-US10624.

XX 05-NOV-1992; 92US-0973337.

PR 05-NOV-1992; 92US-0973337.

XX (SLOK ) SLOAN KETTERING INST CANCER.

PA (SLOK ) SLOAN KETTERING INST CANCER.

XX Fair WR, Heston WDM, Israeli RS;

PI Fair WR, Heston WDM, Israeli RS;

XX WPI; 1994-167129/20.

DR WPI; 1994-167129/20.

XX Prostate-specific membrane antigen and DNA encoding it - is

PT useful for detecting haematogenous micro-metastatic tumour cells

PT useful for detecting haematogenous micro-metastatic tumour cells

PT useful for detecting haematogenous micro-metastatic tumour cells

XX and for identifying ligands which bind to PSM Ag

PS Example ; Page 44; 196pp; English.

XX The inventors attempted to sequence the PSM Ag by modified Edman

CC degradation. Peptides (R55098-107) that gave clear dominant peaks

CC from the digested PSM Ag sample were used. Some of the

CC residues were unidentified, and others were present at very low

CC levels and identified with lower confidence. Some of these peptides

CC were used to design primers to carry out PCR to identify cDNA clones

CC encoding the PSM Ag. R55100 was used to design primers Q65521 and

CC Q65522. A clone, IN-20 was identified as a partial PSM sequence.

CC (The full PSM Ag sequence is shown in R55097 and is encoded

CC by Q65520). The PSM coding sequence is useful for suppressing or

CC modulating the metastatic ability of prostate tumour cells to grow,

CC or for eliminating them. The protein is useful to identify or purify

CC ligands of the Ag. It is also an attractive target for Ab-directed

CC imaging and targeting of prostatic tumour deposits.

XX Sequence 7 AA;

Query Match

Best Local Similarity 44.2%; Score 23; DB 15; Length 7;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DPTFK 5

Db 3 dpmfk 7

RESULT 11

W02239

ID W02239 standard; Peptide; 7 AA.

XX W02239;

AC W02239;

XX 05-NOV-1996 (first entry)

DT 05-NOV-1996 (first entry)

XX Prostate-specific membrane antigen peptide 2T26 3.

DE Prostate-specific membrane antigen; PSM; prostate cancer;

KW metastasis; therapy; diagnosis.

XX Homo sapiens.

OS Homo sapiens.

XX W09626272-A1.

PN W09626272-A1.

XX 29-AUG-1996.

PD 29-AUG-1996.

XX 23-FEB-1996; 96WO-US02424.

PF 23-FEB-1996; 96WO-US02424.

XX 02-JUN-1995; 95US-0470735.

PR 02-JUN-1995; 95US-0470735.

PR 24-FEB-1995; 95US-0394152.

PR 02-JUN-1995; 95US-0466381.  
 XX (SLOK ) SLOAN KETTERING INST CANCER RES.  
 PA Fair WR, Heston WDW, Israeli RS;  
 PI WPI; 1996-402365/40.  
 XX  
 XX DNA encoding alternatively spliced prostate-specific membrane  
 PT antigen - useful to develop prods. for detecting haematogenous  
 PT micrometastatic tumour cells, or prostate cancer progression  
 XX  
 XX Example 1; Page 49; 284pp; English.  
 XX  
 XX Tryptic peptides (W02237-45) were obtd. from the human prostate-  
 CC specific membrane (PSM) antigen (see also W02234) and used to  
 CC design primers (T36795-808) utilised in the isolation of a cDNA  
 CC clone (T36795) coding for PSM; primers A and B (T36795-96) are  
 CC based on amino acids 1-5 of peptide 2726 3 (W02239). The  
 CC isolated cDNA and PSM protein are useful in developing methods for  
 CC the diagnosis and treatment of prostate cancer and metastasis.  
 XX  
 XX Sequence 7 AA;  
 SQ

Query Match 44.2%; Score 23; DB 17; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFK 5  
 DB 3 dpmfk 7

RESULT 12  
 Y55970  
 ID Y55970 standard; Peptide; 9 AA.  
 AC  
 AC Y55970;  
 XX  
 XX 18-FEB-2000 (first entry)  
 DT  
 DE Human SULU1-derived peptide #1.  
 DE  
 KW Antirheumatic; antiarthritic; antiinflammatory; antiallergic; osteopathic;  
 KW antipsoriatic; antiarteriosclerotic; antiasthmatic; immunosuppressive;  
 KW neuroprotective; cardiast; cerebroprotective; cytostatic; antidiabetic;  
 KW vulnery; STE20; protein kinase; STLK3; STLK4; STLK5; STLK6; STLK7;  
 KW ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEEK2, PAK4; antagonist;  
 KW antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;  
 KW inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;  
 KW rhinitis; autoimmunity; organ transplantation; multiple sclerosis;  
 KW myocardial infarction; cardiovascular disease; stroke; renal failure;  
 KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;  
 KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;  
 KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;  
 KW mesangial disorder; growth regulation; wound healing; T cell activation;  
 KW immunosuppressant.  
 XX  
 OS Homo sapiens.  
 XX  
 XX W09553036-A2.  
 PN  
 XX  
 PD 21-OCT-1999.  
 XX  
 XX 13-APR-1999; 99WO-US08150.  
 PF  
 XX  
 XX 14-APR-1998; 98US-0081784.  
 PR  
 XX  
 XX (SUCF-) SUGEN INC.  
 PA  
 XX Plowman G, Martinez R, Whyte D;  
 PI  
 XX

DR WPI; 1999-611301/52.  
 XX  
 XX Novel kinase-related polypeptides used for the diagnosis and treatment  
 PT of kinase-related diseases and disorders  
 XX  
 XX Disclosure; Page 379; 387pp; English.  
 XX  
 XX This sequence represents a peptide fragment from a novel STE20-related  
 CC protein kinases. The invention relates to nucleic acid molecules encoding  
 CC a kinase polypeptide selected from STLK2, STLK3, STLK4, STLK5, STLK6,  
 CC STLK7, ZC1, ZC2, ZC3, ZC4, KHS2, SULU1, SULU3, GEEK2, PAK4 and PAK5. The  
 CC proteins are used to identify agonists and antagonists, and to raise  
 CC antibodies. The polynucleotides are useful in gene therapy protocols. The  
 CC polynucleotides, polypeptides, antibodies, antagonists and agonists may  
 CC be used to treat diseases such as immune-related disorders and diseases  
 CC (e.g. rheumatoid arthritis, artherosclerosis, chronic inflammatory bowel  
 CC disease (e.g. Crohn's disease), asthma, osteoarthritis, psoriasis,  
 CC atherosclerosis, rhinitis, autoimmunity, and organ transplantation,  
 CC chronic inflammatory pelvic disease, multiple sclerosis, organ  
 CC transplantation, myocardial infarction, cardiovascular disease, stroke,  
 CC renal failure, oxidative stress-related neurodegenerative disorders (e.g.  
 CC amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome),  
 CC cancer, cardiomyopathies, ischemic disorders, inflammatory disorders,  
 CC diabetes mellitus, fibrotic and mesangial disorders. The proteins may  
 CC also be useful for cell growth regulation (e.g. in wound healing), T cell  
 CC activation, mitosis control, and as immunosuppressants.  
 XX  
 XX Sequence 9 AA;  
 SQ

Query Match 44.2%; Score 23; DB 20; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KENYR 9  
 DB 5 kedyr 9

RESULT 13  
 R24699  
 ID R24699 standard; peptide; 6 AA.  
 AC  
 AC R24699;  
 XX  
 XX 03-DEC-1992 (first entry)  
 DT  
 XX Immunomodulatory peptide.  
 DE  
 XX Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;  
 KW inflammation; wounds; lymphocyte; vaccine.  
 KW  
 XX Synthetic.  
 OS  
 XX W09209628-A.  
 PN  
 XX 11-JUN-1992.  
 PD  
 XX 22-NOV-1991; 91WO-US08795.  
 PF  
 XX 23-NOV-1990; 90US-0617494.  
 PR  
 XX (IMMU-) IMMUNODYNAMICS INC.  
 PA  
 XX Atkin A;  
 PI  
 XX WPI; 1992-217021/26.  
 DR  
 XX New synthetic immunomodulatory peptide(s) - for treating  
 PT immunodeficiencies, immunosuppression and T-cell subset  
 PT deviations and immuno-therapy of infections, inflammation, wounds  
 PT etc.  
 XX

PS Claim 10; Page 36; 52pp; English.

CC The immunomodulatory peptide is a specific example of a peptide cpd.  
 CC (or an acid or base salt) constructed by combination and/or  
 CC overlapping of the amino acid sequences AIBXB2A2, A3B3XA4B4,  
 CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, and B11B12 (X= Ala, Gly,  
 CC Ile, Leu, Phe or Val, A1-A12 each= Arg, Asn, Gln, Lys, Phe or Val;  
 CC B1-B12 each= Asp, Glu, Tyr, Phe or Val. The synthetic peptide may  
 CC be used for immunomodulation of various immunodeficiencies and  
 CC immunosuppressed conditions. T-cell subset and lymphocyte deviations,  
 CC enhancement of a vaccine efficacy, as well as for immunotherapy,  
 CC including infections, local or systemic complications of non-  
 CC infectious diseases, postoperative inflammations, wounds and burns.  
 CC See also R24583-R24701.

XX Sequence 6 AA;

Query Match 42.3%; Score 22; DB 13; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+05;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 FKENVYR 9  
 :|| ||  
 Db 1 ykelyr 6

RESULT 14

R27999  
 ID R27999 standard; peptide; 8 AA.

XX

AC R27999;

XX

DT 17-MAR-1993 (first entry)

XX Tryptic peptide fragment #17.

DE Insulin receptor substrate-1; IRS-1; antibody; proteolytic cleavage;

KW 1D-SDS PAGE.

XX Rattus rattus.

OS W09213083-A.

XX 06-AUG-1992.

PF 17-JAN-1992; 92WO-US00437.

XX 18-JAN-1991; 91US-0643982.

PR (JOSL-) JOSLIN DIABETES CENT INC.

XX Kahn CR, Rothenberg PL, White MF;

PI WPI; 1992-365881/44.

DR Purified nucleic acid encoding Insulin Receptor Substrate - used

XX to prepare IRS-1, for diagnosis and treatment of insulin related

CC diseases and abnormal cellular proliferation

XX Disclosure; Page 26; 128pp; English.

XX The sequences given in R27983-8000 and R28044 are fragments from

CC insulin receptor substrate-1 (IRS-1). These fragments were used to

CC determine the sequence of IRS-1 and to distinguish it from proteins

CC which are co-purified with it. Antibodies were raised against the

CC IRS-1 proteins and were used to remove them from the reaction media.

CC These peptides were formed by proteolytic cleavage of proteins

XX isolated by ID-SDS PAGE to be approx. 185 kD.

XX Sequence 8 AA;

Query Match 42.3%; Score 22; DB 13; Length 8;  
 Best Local Similarity 42.9%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TFKENYR 9  
 ||| ||:  
 Db 1 tfeesfq 7

RESULT 15

Y37952  
 ID Y37952 standard; Peptide; 9 AA.

XX

AC Y37952;

XX

DT 29-SEP-1999 (first entry)

XX

DE Human CERB2 oncogene-derived HLA-binding peptide.

XX

KW Immunogen; HLA; human leukocyte antigen; binding motif; antiviral;

KW MHC; major histocompatibility complex; viral infection; anticancer;

KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.

XX Homo sapiens.

OS W09403205-A1.

XX

PD 17-FEB-1994.

XX

PF 06-AUG-1993; 93WO-US07421.

XX

PR 05-MAR-1993; 93US-0027746.

PR 07-AUG-1992; 92US-0926666.

XX (CYTE-) CYTEL CORP.

PA

XX

PI Celis E, Grey HM, Kubo RT, Sette A;

XX WPI; 1994-065403/08.

DR

XX

PT Peptide which specifically binds selected MHC allele - used to

CC induce an immune response for treatment or prevention of viral

CC infection or cancer, or for diagnosis

XX Disclosure; Page 103; 150pp; English.

XX

CC The sequence is a specific example of a group of new immunogenic

CC peptides having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding

CC motif. For example, the peptides having an HLA-A3.2 binding motif

CC each have 9-10 residues and contain, from the N-terminus to the

CC C-terminus, (a) a first conserved residue selected from L, M, I,

CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of

CC K, R, Y, H or F, where the first and second conserved residues are

CC separated by 6-7 residues. The peptides are capable of binding

CC selected MHC molecules and inducing an immune response. They can be

CC used to treat and/or prevent viral infection and cancer, e.g. prostate

CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce

CC antibodies for use as diagnostic or therapeutic agents. The peptides

CC can also be used as diagnostic agents.

XX Sequence 9 AA;

Query Match 42.3%; Score 22; DB 15; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2.1e+05;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 FKENVY 8  
 ||: ||  
 Db 5 fedny 9

Search completed: January 17, 2001, 13:42:25  
Job time: 139 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:43:38 ; Search time 36.59 Seconds  
(without alignments)  
16.701 Million cell updates/sec

Title: US-08-765-837-9

Perfect score: 52

Sequence: 1 DPTFKENYR 9

Scoring table: BLOSUMP62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 787

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	42.3	8	2 S21273	cellulase (EC 3.2.1.4)
2	19	36.5	8	2 JS0317	leucokinin VII - M
3	19	36.5	9	2 PT0324	Ig heavy chain CRD
4	18	34.6	4	2 S53508	starvation-induced
5	18	34.6	8	2 S71919	alcohol dehydrogen
6	17	32.7	8	2 S08996	hypertrehalosemic
7	17	32.7	8	2 B49823	adipokinetic hormo
8	17	32.7	8	2 B49600	neuropeptide led-c
9	17	32.7	8	2 A61597	cytochrome P450 AL
10	16	30.8	8	2 S70727	199F protein - Shl
11	16	30.8	9	2 A42266	peptidylglycine mo
12	15	28.8	6	2 S11556	hydrogensulfite re
13	15	28.8	8	2 S43971	tumor-associated a
14	15	28.8	8	2 S43972	calsequestrin, car
15	15	28.8	9	2 A61230	Ig heavy chain CRD
16	15	28.8	9	2 PT0270	Ig heavy chain CRD
17	15	28.8	8	2 PT0285	Ig heavy chain CRD
18	14	26.9	8	2 PT0030	inulinase (EC 3.2.
19	14	26.9	9	2 A60108	exotoxin A - Strep
20	13	25.0	4	2 I51049	metallothionein-A
21	13	25.0	5	2 B60274	major protein anti
22	13	25.0	5	2 PT0281	Ig heavy chain CRD
23	13	25.0	6	2 A19780	transferrin - bovi
24	13	25.0	6	2 PT0709	T-cell receptor be
25	13	25.0	7	2 S71867	glutathione trans
26	13	25.0	7	2 A38081	amine oxidase (cop
27	13	25.0	7	4 I56695	hypothetical L2 pr
28	13	25.0	8	2 S08995	hypertrehalosemic
29	13	25.0	8	2 A49823	adipokinetic hormo

30 13 25.0 8 2 A44960 neuropeptide led-c  
31 13 25.0 8 2 S15422 adipokinetic hormo  
32 13 25.0 8 2 A43976 hypertrehalosemic  
33 13 25.0 8 2 B43976 hypertrehalosemic  
34 13 25.0 8 2 S11545 adipokinetic hormo  
35 13 25.0 8 2 A58641 adipokinetic hormo  
36 13 25.0 8 2 S63493 dissimilatory sulf  
37 13 25.0 8 2 A05169 neuropeptide M-r -  
38 13 25.0 8 2 S21663 neuropeptide - flo  
39 13 25.0 8 2 I64832 Ca2+-transporting  
40 13 25.0 9 2 B45796 dihydroliposamide S  
41 13 25.0 9 2 A24244 adipokinetic hormo  
42 13 25.0 9 2 S30494 cat gene leader pe  
43 13 25.0 9 2 B24362 chloramphenicol O-  
44 13 25.0 9 2 S66607 quinaldine 2-oxidor  
45 13 25.0 9 2 C36730 hutU protein kle

#### ALIGNMENTS

RESULT 1

S21273 cellulase (EC 3.2.1.4) - Clostridium thermoCELLUM (fragment)  
N: Alternate names: endo-1,4-beta-glucanase  
C: Species: Clostridium thermoCELLUM  
C: Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 22 N. 7-1996  
C: Accession: S21273  
R: Romanec, M.P.M.; Fauth, U.; Kobayashi, T.; Huskisson, N.S.; Barker, P.J.; Demain, Biochem. J. 283, 69-73, 1992  
A: Title: Purification and characterization of a new endoglucanase from Clostridium th  
A: Reference number: S21273; MUID: 92231850  
A: Accession: S21273  
A: Molecule type: protein  
A: Residues: 1-8 <ROM>  
C: Function:  
A: Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as  
A: Pathway: cellulose degradation  
C: Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 42.3%; Score 22; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PTFKE 6

Db 2 PTFKE 6

RESULT 2

JS0317 leucokinin VII - Madeira cockroach  
C: Species: Leucophaea madeirae (Madeira cockroach)  
C: Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Jun-2000  
C: Accession: JS0317  
R: Holman, G.M.; Cook, B.J.; Nachman, R.J.  
Comp. Biochem. Physiol. C 88, 31-34, 1987  
A: Title: Isolation, primary structure and synthesis of leucokinin VII and VIII: the  
A: Reference number: JS0317  
A: Accession: JS0317  
A: Molecule type: protein  
A: Residues: 1-8 <HOL>  
C: Comment: Leucokinin, a family of cephalomyotropic peptides, stimulate contractile  
C: Keywords: amidated carboxyl end; cephalomyotropic peptide  
F: 8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 36.5%; Score 19; DB 2; Length 8;

Best Local Similarity 75.0%; Pred. No. 2e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFK 4

Db 1 DPAF 4  
|| |

## RESULT 3

PT0324  
Ig heavy chain CRD3 region (clone J2-106C) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0324  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J  
A:Reference number: PT0222; MUID:91108337  
A:Accession: PT0324  
A:Molecule type: DNA  
A:Residues: 1-9 <YAM>  
A:Experimental source: B lymphocyte  
A:Keywords: heterotetramer; immunoglobulin

Query Match 36.5%; Score 19; DB 2; Length 9;  
Best Local Similarity 42.9%; Pred. No. 2e+05;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PTFKENY 8  
| : | : |  
Db 2 PGTGESY 8

## RESULT 4

S53508  
starvation-induced ribonuclease - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 07-May-1999  
C:Accession: S53508  
R:Koeck, M.; Loeffler, A.; Abel, S.; Glund, K.  
Plant Mol. Biol. 27, 477-485, 1995  
A:Title: cDNA structure and regulatory properties of a family of starvation-induced ribo  
A:Reference number: S53506; MUID:95201242  
A:Accession: S53508  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-4 <KOE>

Query Match 34.6%; Score 18; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PTF 4  
|| |  
Db 2 PTF 4

## RESULT 5

S71919  
alcohol dehydrogenase (EC 1.1.1.1) - grass carp (fragment)  
C:Species: Ctenopharyngodon idella (grass carp)  
C:Date: 14-Apr-1998 #sequence\_revision 24-Apr-1998 #text\_change 07-May-1999  
C:Accession: S71919  
R:Tsuji, H.T.; Mock, W.Y.; Lau, K.K.; Fong, W.P.  
Biochim. Biophys. Acta 1296, 41-46, 1996  
A:Title: Proteolytic activation of grass carp (Ctenopharyngodon idellus) liver alcohol de  
A:Reference number: S71919; MUID:96350418  
A:Accession: S71919  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <TSU>  
A:Note: the source is designated Ctenopharyngodon idellus  
C:Keywords: NAD; oxidoreductase

Query Match 34.6%; Score 18; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPT 3  
|| |  
Db 2 DPT 4

## RESULT 6

S08996  
hypertrehalosemic hormone II - oriental cockroach  
N:Alternate names: Pea-CAH-II  
C:Species: Blatta orientalis (oriental cockroach)  
C:Date: 30-Jun-1992 #sequence\_revision 24-Oct-1997 #text\_change 31-Oct-1997  
C:Accession: S08996  
R:Gaede, G.; Rinehart, K.L.  
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990  
A:Title: Primary structures of hypertrehalosemic neuropeptides isolated from the cor  
entails and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bomb  
A:Reference number: S08995; MUID:90253659  
A:Accession: S08996  
A:Molecule type: protein  
A:Residues: 1-8 <GAE>  
A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we h  
C:Superfamily: adipokinetic hormone  
C:Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TFKENY 8  
|| |  
Db 3 TFTPNN 8

## RESULT 7

B49823  
adipokinetic hormone II - American cockroach  
N:Alternate names: neuropeptide M-II; periplanetin CC-1  
C:Species: Periplaneta americana (American cockroach)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 07-May-1999  
C:Accession: B49823; A05170  
R:Scarborough, R.W.; Jamieson, G.C.; Kalish, F.; Kramer, S.J.; McEnroe, G.A.; Miller,  
Proc. Natl. Acad. Sci. U.S.A. 81, 5575-5579, 1984  
A:Title: Isolation and primary structure of two peptides with cardioacceleratory and  
A:Reference number: A49823; MUID:84298179  
A:Accession: B49823  
A:Molecule type: protein  
A:Residues: 1-8 <SCA>  
R:Witten, J.L.; Schaffer, M.H.; O'Shea, M.; Cook, J.C.; Hemling, M.E.; Rinehart Jr.,  
Biochem. Biophys. Res. Commun. 124, 350-358, 1984  
A:Title: Structures of two cockroach neuropeptides assigned by fast atom bombardment  
A:Reference number: A90118; MUID:85046530  
A:Accession: A05170  
A:Molecule type: protein  
A:Residues: 'E', 2-8 <WIT>  
C:Superfamily: adipokinetic hormone  
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyrogluta  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 32.7%; Score 17; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TFKENY 8  
|| |

Db 3 TFTPWW 8

## RESULT 8

B44960  
 neuropeptide Led-CC-II - Colorado potato beetle  
 C:Species: Leptinotarsa decemlineata (Colorado potato beetle)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
 C:Accession: B44960  
 R:Gaede, G.; Kellner, R.  
 Peptides 10, 1287-1289, 1989  
 A:Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and  
 A:Reference number: A44960; MUID:90160053  
 A:Accession: B44960  
 A:Molecule type: protein  
 A:Residues: 1-8 <GAE>  
 C:Superfamily: adipo-kinetic hormone  
 C:Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TFKENY 8

|||

Db 3 TFTPWW 8

## RESULT 9

A61597  
 cytochrome P450 AL-1 - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
 C:Accession: A61597  
 R:Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.  
 Drug Metab. Dispos. 19, 291-297, 1991  
 A:Title: Purification and aminopyrine monooxygenase activity of liver microsomal cytochr  
 A:Reference number: A61597; MUID:91292910  
 A:Accession: A61597  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-8 <SHI>

Query Match 32.7%; Score 17; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PTFKEN 7

|||

Db 3 PXXVES 8

## RESULT 10

ipgF protein - Shigella flexneri (fragment)  
 C:Species: Shigella flexneri  
 C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999  
 C:Accession: S70727  
 R:Aliaoui, A.; Sansonetti, P.J.; Menard, R.; Barzu, S.; Mounier, J.; Phalipon, A.; Parsc  
 Mol. Microbiol. 17, 461-470, 1995  
 A:Title: MxiG, a membrane protein required for secretion of Shigella spp. Ipa invasins:  
 A:Reference number: S70727; MUID:96100445  
 A:Accession: S70727  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-8 <ALL>

A:Cross-references: EMBL:248957; NID:g929880; PIDN:CAA88921.1; PID:g929881  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995  
 C:Genetics:

A:Gene: ipgF

Query Match 30.8%; Score 16; DB 2; Length 8;  
 Best Local Similarity 60.0%; Pred. No. 2e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 KENYR 9

|||

Db 2 KONNR 6

## RESULT 11

A42266  
 peptidylglycine monooxygenase (EC 1.14.17.3), rPAM-5 - rat (fragment)  
 N:Alternate names: peptidylglycine alpha-amidating monooxygenase  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-May 1995  
 C:Accession: A42266  
 R:Elipper, B.A.; Green, C.B.; Campbell, T.A.; Stoffers, D.A.; Keutmann, H.T.; Mains, R  
 J. Biol. Chem. 267, 4008-4015, 1992  
 A:Title: Alternative splicing and endoproteolytic processing generate tissue-specific  
 A:Reference number: A42266; MUID:92156145  
 A:Accession: A42266  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-9 <EIP>  
 A:Experimental source: pituitary  
 A:Note: sequence extracted from NCBI backbone (NCBIN:82733, NCBI:P42750)  
 C:Keywords: oxidoreductase

Query Match 30.8%; Score 16; DB 2; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 2e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 FKENY 8

|||

Db 5 FKDTF 9

## RESULT 12

S11556  
 hydrogensulfite reductase (EC 1.8.99.3) chain 2 - Desulfovibrio thermophilus (fragmen  
 N:Alternate names: bisulfite reductase; desulfofusicidin  
 C:Species: Desulfovibrio thermophilus  
 C:Date: 19-Mar-1997 #sequence\_revision 12-Dec-1997 #text\_change 30-Jan-1998  
 C:Accession: S11556  
 R:Fauque, G.; Lino, A.R.; Czechowski, M.; Kang, L.; Dervartanian, D.V.; Moura, J.J.G.  
 Biochim. Biophys. Acta 1040, 112-118, 1990  
 A:Title: Purification and characterization of bisulfite reductase (desulfofusicidin) f  
 A:Reference number: S11024; MUID:90335276  
 A:Accession: S11556  
 A:Molecule type: protein  
 A:Residues: 1-6 <FAU>  
 C:Keywords: oxidoreductase

Query Match 28.8%; Score 15; DB 2; Length 6;  
 Best Local Similarity 40.0%; Pred. No. 2e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 KENYR 9

|||

Db 2 EEKYK 6

## RESULT 13

S43971  
 tumor-associated antigen MMT1 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Oct-1994 #sequence\_revision 17-Nov-1995 #text\_change 18-Aug-2000  
 C:Accession: S43971

R:Mandelboim, O.; Berke, G.; Fridkin, M.; Feldman, M.; Eisenstein, M.; Eisenbach, L.  
Nature 369, 67-71, 1994  
A:Title: CTL induction by a tumour-associated antigen octapeptide derived from a murine  
A:Reference number: S43971; MUID:94217811  
A:Accession: S43971  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <MAN>  
C:Superfamily: unassigned animal peptides

Query Match 28.8%; Score 15; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 FKFN 7  
|::|  
Db 1 FEQN 4

RESULT 14  
S43972  
tumour-associated antigen MUT2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Oct-1994 #sequence\_revision 17-Nov-1995 #text\_change 18-Aug-2000  
C:Accession: S43972  
R:Mandelboim, O.; Berke, G.; Fridkin, M.; Feldman, M.; Eisenstein, M.; Eisenbach, L.  
Nature 369, 67-71, 1994  
A:Title: CTL induction by a tumour-associated antigen octapeptide derived from a murine  
A:Reference number: S43971; MUID:94217811  
A:Accession: S43972  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <MAN>  
C:Superfamily: unassigned animal peptides

Query Match 28.8%; Score 15; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 FKFN 7  
|::|  
Db 1 FEQN 4

RESULT 15  
A61230  
calsequestrin, cardiac and slow skeletal muscle - northern leopard frog (fragment)  
N:Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-  
C:Species: Rana pipiens (northern leopard frog)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 16-Jul-1999  
C:Accession: A61230  
R:McLeod, A.G.; Shen, A.C.Y.; Campbell, K.P.; Michalak, M.; Jorgensen, A.O.  
Circ. Res. 69, 344-359, 1991  
A:Title: Frog cardiac calsequestrin. Identification, characterization, and subcellular  
rdium.  
A:Reference number: A61230; MUID:91316784  
A:Accession: A61230  
A:Molecule type: protein  
A:Residues: 1-9 <MCL>  
C:Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protein  
C:Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to  
C:Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in vi  
C:Superfamily: calsequestrin  
C:Keywords: calcium binding; cardiac muscle; glycoprotein; heart; phosphoprotein; skelet

Query Match 28.8%; Score 15; DB 2; Length 9;  
Best Local Similarity 66.7%; Pred. No. 2e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PTF 4

Db |::|  
7 PTY 9

Search completed: January 17, 2001, 13:43:38  
Job time: 157 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 17, 2001, 13:47:10 ; Search time 20.9 Seconds  
(without alignments)  
13.907 Million cell updates/sec

Title: US-08-765-837-9  
Perfect score: 52  
Sequence: 1 DPTFENYR 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues  
Total number of hits satisfying chosen parameters: 213

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	36.5	8	1 FARL_PANRE	P41872 panagrellus
2	19	36.5	8	1 LCK1_LEUMA	P21140 leucophaea
3	19	36.5	8	1 LCK7_PANRE	P19989 leucophaea
4	19	36.5	9	1 FAR2_PANRE	P41873 panagrellus
5	18	34.6	9	1 FAR5_ASCSU	P43170 ascaris suu
6	17	32.7	8	1 HTF2_PERAM	P04549 periplaneta
7	17	32.7	8	1 LCK2_LEUMA	P21141 leucophaea
8	14	26.9	9	1 ULAD_HUMAN	P31929 homo sapien
9	13	25.0	6	1 TMOF_SARBU	P41495 sarcophaga
10	13	25.0	7	1 ALL4_CARMA	P81807 carcinus ma
11	13	25.0	7	1 FAR1_HELTI	P41871 helisoma tr
12	13	25.0	7	1 FAR4_PANRE	P41875 panagrellus
13	13	25.0	8	1 AKH_MELML	P25423 melolontha
14	13	25.0	8	1 B44K_PORGI	P81886 porphyromon
15	13	25.0	8	1 HTF1_PERAM	P04548 periplaneta
16	13	25.0	8	1 HTF_TENMO	P25419 tenebrio mo
17	13	25.0	9	1 FLA2_TREHY	P80159 treponema h
18	13	25.0	9	1 HUTA_KLEAE	P2381 klebsiella
19	13	25.0	9	1 LPCA_STAAU	P36884 staphylococ
20	12	23.1	5	1 PRCT_PERAM	P01373 periplaneta
21	12	23.1	6	1 CIP1_MYTED	P13736 mytilus edu
22	12	23.1	6	1 CIP2_MYTED	P13737 mytilus edu
23	12	23.1	7	1 FAR2_ASCSU	P31890 ascaris suu
24	12	23.1	8	1 ACI_THUAL	P18691 thunnus alb
25	12	23.1	8	1 FAR7_ASCSU	P43171 ascaris suu
26	12	23.1	8	1 LCK4_LEUMA	P21143 leucophaea
27	12	23.1	8	1 RS10_SALTY	O68928 salmonella
28	12	23.1	9	1 COXE_THUOB	P80975 thunnus obe
29	12	23.1	9	1 OXYT_EISFO	P42998 eisenia foe
30	12	23.1	9	1 RS10_SERNA	O68936 serratia ma
31	12	23.1	9	1 TKC1_CALVO	P41517 calliphora
32	12	23.1	9	1 UP33_HUMAN	P30089 homo sapien
33	11	21.2	7	1 FAR6_CALVO	P41866 calliphora

34	11	21.2	7	1 WWA3_ACHFU	P35921 actulina fu
35	11	21.2	8	1 AKH_TABAT	P14595 t. t. mus atr
36	11	21.2	8	1 CLP_THICU	P80488 t. t. obacillu
37	11	21.2	8	1 RS1_ERWCH	P37985 erwinia chr
38	11	21.2	8	1 UC26_MAIZE	P80632 zea mays (m
39	11	21.2	8	1 UPAA_HUMAN	P30096 homo sapien
40	11	21.2	9	1 MOSF_CLYJA	P19853 clypeaster
41	11	21.2	9	1 NEUX_HUMAN	P04277 homo sapien
42	11	21.2	9	1 NEUX_RAT	P11382 rattus norv
43	11	21.2	9	1 XYL4_STRSQ	P19149 streptomyce
44	10	19.2	6	1 VP19_HSVIK	P23210 herpes simp
45	10	19.2	7	1 ALL2_CARMA	P81805 carcinus ma

ALIGNMENTS

RESULT 1					
FARL_PANRE					
ID FARL_PANRE	STANDARD;	PRT;	8 AA.		
AC P41872;					
DT 01-NOV-1995	(Rel. 32, Created)				
DT 01-NOV-1995	(Rel. 32, Last sequence update)				
DT 01-NOV-1995	(Rel. 32, Last annotation update)				
DE FMRFAMIDE-LIKE NEUROPEPTIDE PFL (SDPNELRF-AMIDE).					
OS Panagrellus redivivus.					
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;					
OC Panagrolaimidae; Panagrolaimidae; Panagrellus.					
RN [1]					
RP SEQUENCE.					
RX MEDLINE=93027659; PubMed=1408999;					
RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,					
RA Garrison R.D., Williams J.F., Friedman A.R.;					
RT "Two FMRFamide-like peptides from the free-living nematode					
RT Panagrellus redivivus.";					
RL Peptides 13:209-214(1992).					
CC -!- FUNCTION: MYOACTIVE.					
CC -!- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS.					
CC -!- CAUDALLY TO THE BASE OF THE PHARYNX.					
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)					
CC FAMILY.					
KW Neuropeptide; Amidation.					
FT MOD_RES 8 8	AMIDATION.				
SQ SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;					
Query Match	36.5%;	Score 19;	DB 1;	Length 8;	
Best Local Similarity	75.0%;	Pred. No. 8.9e+04;			
Matches 3;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
Qy 1 DPTF 4					
Db 2 DPNF 5					
RESULT 2					
LCK1_LEUMA					
ID LCK1_LEUMA	STANDARD;	PRT;	8 AA.		
AC P21140;					
DT 01-MAY-1991	(Rel. 18, Created)				
DT 01-MAY-1991	(Rel. 18, Last sequence update)				
DT 01-MAY-1991	(Rel. 18, Last annotation update)				
DE LEUCOKININ I (L-I).					
OS Leucophaea maderae (Madeira cockroach).					
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;					
OC Blaberoidea; Blaberidae; Leucophaea.					
RN [1]					
RP SEQUENCE, AND SYNTHESIS.					
RC TISSUE-HEAD;					
RA Holman G.M., Cook B.J., Nachman R.J.;					
RT "Isolation, primary structure and synthesis of two neuropeptides					
RT from Leucophaea maderae: members of a new family of					

RT Cephalomyotroptins.";  
 RL Comp. Biochem. Physiol. 84C:205-211(1986).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC -!- ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 KW Neuropeptide; Amidation.  
 FT MOD.RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 893 MW; DC6365B449CDC76A CRC64;

Query Match 36.5%; Score 19; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTF 4  
 |||  
 Db 1 DPAF 4

RESULT 3  
 LCK7\_LEUMA STANDARD; PRT; 8 AA.  
 AC P1989;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE LEUCOKININ VII (L-VII).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidea; Blaberidae; Leucophaea.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-HEAD;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of leucokinins VII and  
 RT VIII: the final members of this new family of cephalomyotrophic  
 RT peptides isolated from head extracts of Leucophaea maderae.";  
 RL Comp. Biochem. Physiol. 88C:31-34(1987).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 DR PIR; JS0317; JS0317.  
 KW Neuropeptide; Amidation.  
 FT MOD.RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 866 MW; DC6365A5B9CDC76A CRC64;

Query Match 36.5%; Score 19; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTF 4  
 |||  
 Db 1 DPAF 4

RESULT 4  
 FAR2\_PANRE STANDARD; PRT; 9 AA.  
 AC P41873;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE PF2 (SADPNFLRF-AMIDE).  
 OS Panagrellus redivivus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93027659; PubMed=1408999;  
 RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,  
 RA Garrison R.D., Williams J.F., Friedman A.R.;

RT "Two FMRFamide-like peptides from the free-living nematode  
 RT Panagrellus redivivus.";  
 RL Peptides 13:209-214(1992).  
 CC -!- FUNCTION: MYOACTIVE.  
 CC -!- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS LOCATED  
 CC CAUDALLY TO THE BASE OF THE PHARYNX.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD.RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 1066 MW; DA0B0729C4576AAD CRC64;

Query Match 36.5%; Score 19; DB 1; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTF 4  
 |||  
 Db 3 DPNF 6

RESULT 5  
 FAR5\_ASCSU STANDARD; PRT; 9 AA.  
 ID FAR5\_ASCSU  
 AC P43170;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE AF5.  
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;  
 OC Ascarididae; Ascaris.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95380362; PubMed=7651904;  
 RA Cowden C., Stretton A.O.W.;  
 RT "Eight novel FMRFamide-like neuropeptides isolated from the nematode  
 RT Ascaris suum.";  
 RL Peptides 16:491-500(1995).  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD.RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 1052 MW; 340B0059D1B76338 CRC64;

Query Match 34.6%; Score 18; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PTF 4  
 |||  
 Db 4 PTF 6

RESULT 6  
 HTF2\_PERAM STANDARD; PRT; 8 AA.  
 ID HTF2\_PERAM  
 AC P04549;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE HYPERTREHALOSAEMIC FACTOR II (NEUROPEPTIDE M-II) (PERIPLANETIN CC-2)  
 DE (PEA-CAH-II) (LED-CC-II) (HYPERTREHALOSAEMIC NEUROPEPTIDE II).  
 OS Periplaneta americana (American cockroach),  
 OS Leptinotarsa decemlineata (Colorado potato beetle), and  
 OS Blattia orientalis (Oriental cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blattioidea; Blattidae; Periplaneta.  
 RN [1]  
 RP SEQUENCE.

RC SPECIES=P-AMERICANA;  
 RX MEDLINE=85046530; PubMed=6548628;  
 RA Witten J.L., Schaffer M.H., O'Shea M., Cook J.C., Hemling M.E.,  
 RA Rinehart K.L. Jr.;  
 RT "Structures of two cockroach neuropeptides assigned by fast atom  
 RT bombardment mass spectrometry.";  
 RL Biochem. Biophys. Res. Commun. 124:350-358(1984).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=P-AMERICANA;  
 RX MEDLINE=84298179; PubMed=6591205;  
 RA Scarborough R.M., Jamieson G.C., Kalish F., Kramer S.J., McEnroe G.A.,  
 RA Miller C.A., Schooley D.A.;  
 RT "Isolation and primary structure of two peptides with  
 RT cardioacceleratory and hyperglycemic activity from the corpora  
 RT cardiaca of *Periplaneta americana*.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).  
 RN [3]  
 RP SEQUENCE.  
 RC SPECIES=L-DECEMLINEATA; TISSUE=CORPORA CARDIACA;  
 RX MEDLINE=90160053; PubMed=2576128;  
 RA Gaede G., Kellner R.;  
 RT "The metabolic neuropeptides of the corpus cardiacum from the potato  
 RT beetle and the American cockroach are identical.";  
 RL Peptides 10:1287-1289(1989).  
 RN [4]  
 RP SEQUENCE.  
 RC SPECIES=B-ORIENTALIS; TISSUE=CORPORA CARDIACA;  
 RX MEDLINE=90253659; PubMed=2340112;  
 RA Gaede G., Rinehart K.L. Jr.;  
 RT "Primary structures of hypertrehalosaemic neuropeptides isolated from  
 RT the corpora cardiaca of the cockroaches *Leucophaea maderae*,  
 RT *Gromphadorhina portentosa*, *Blattella germanica* and *Blatta orientalis*  
 RT and of the stick insect *Extatosoma tiaratum* assigned by tandem fast  
 RT atom bombardment mass spectrometry.";  
 RL Biol. Chem. Hoppe-Seyler 371:345-354(1990).  
 RC CC  
 CC -1- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT  
 CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS  
 CC THE MAJOR CARBOHYDRATE IN THE AKH / HRTH / RPCH FAMILY).  
 CC -1- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.  
 DR PIR; A05170; A05170.  
 DR PIR; S08996; S08996.  
 DR PIR; B44960; B44960.  
 DR PIR; B49823; B49823.  
 DR INTERPRO: IPR002047;  
 DR PROSITE; PS00256; AKH; 1.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 1006 MW; 86745771A9D1A736 CRC64;

Query Match 32.7%; Score 17; DB 1; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TFKENY 8  
 DB 3 TFTPWW 8

RESULT 7  
 LCK2\_LEUMA  
 ID LCK2\_LEUMA STANDARD; PRT; 8 AA.  
 AC P21141;  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DE LEUCOKININ II (L-II).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 OC Blaberoidea; Blaberidae; Leucophaea.

RN [1]  
 RP SEQUENCE AND SYNTHESIS.  
 RC TISSUE=HEAD;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of two neuropeptides  
 RT from *Leucophaea maderae*: members of a new family of  
 RT cephalomyotropins.";  
 RL Comp. Biochem. Physiol. 84C:205-211(1986).  
 CC -1- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -1- SIMILARITY: TO THE OTHER LEUCOKININS.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 852 MW; DC6365A5B9C8676A CRC64;

Query Match 32.7%; Score 17; DB 1; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 8.9e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTF 4  
 DB 1 DPGF 4

RESULT 8  
 ULAD\_HUMAN  
 ID ULAD\_HUMAN STANDARD; PRT; 9 AA.  
 AC P31929;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF LIVER TISSUE (SPOT 106) (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE  
 RC TISSUE=LIVER;  
 RX MEDLINE=94147969; PubMed=8313870;  
 RA Hughes G.J., Frutiger S., Paquet N., Pasquali C., Sanchez J.-C.,  
 RA Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;  
 RT "Human liver protein map: update 1993.";  
 RL Electrophoresis 14:1216-1222(1993).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6, ITS MW IS: 15 KDA.  
 DR SWISS-2DPAGE; P31929; HUMAN.  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1129 MW; D02DFB41B6D33322 CRC64;

Query Match 26.9%; Score 14; DB 1; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 KENY 8  
 DB 4 KQTY 7

RESULT 9  
 TMOF\_SARBU  
 ID TMOF\_SARBU STANDARD; PRT; 6 AA.  
 AC P41495;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE TRYPsin-MODULATING OOSTATIC FACTOR (TMOF).  
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Sarcophagidae; Sarcophaga.  
 RN [1]

## RP SEQUENCE, AND SYNTHESIS.

RC TISSUE-OVARY;  
 RX MEDLINE=94211930; PubMed=8159807;  
 RA Bylemans D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,  
 de Loof A.;  
 RT "Sequencing and characterization of trypsin modulating oostatic  
 factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria  
 RT (Sarcophaga) bullata.";  
 RL Regul. Pept. 50:61-72(1994).  
 CC -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS  
 CC IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN  
 CC CONCENTRATION IN THE HEMOLYMPH RESULTING IN INHIBITION OF OOCYTE  
 CC DEVELOPMENT.  
 CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR  
 CC EPITHELIUM AFTER A BLOOD MEAL.  
 KW Hormone.  
 SQ SEQUENCE 6 AA; 695 MW; 61E72451B7642000 CRC64;

Query Match 25.0%; Score 13; DB 1; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPT 3  
 :||  
 Db 1 NPT 3

## RESULT 10

ALL4\_CARMA  
 ID ALL4\_CARMA STANDARD; PRT; 7 AA.  
 AC P81807;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINOSTATIN 4.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 CC Eubrachiura; Portunoidae; Portunidae; Carcinus.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 SQ SEQUENCE 7 AA; 782 MW; 672879DCB476AC0 CRC64;

Query Match 25.0%; Score 13; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DP 2  
 :||  
 Db 1 DP 2

## RESULT 11

FAR1\_HELTI  
 ID FAR1\_HELTI STANDARD; PRT; 7 AA.  
 AC P41871;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE GDPFLRF-AMIDE.  
 OS Helisoma trivolvis (Snail).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 OC Planorbidae; Helisoma.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=KIDNEY;  
 RX MEDLINE=94286417; PubMed=7912428;  
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
 RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
 RT trivolvis.";  
 RL Peptides 15:31-36(1994).  
 CC -!- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING  
 CC THE KIDNEY, MANTLE AND SKIN.  
 CC -!- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;

Query Match 25.0%; Score 13; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DP 2  
 :||  
 Db 2 DP 3

## RESULT 12

FAR4\_PANRE  
 ID FAR4\_PANRE STANDARD; PRT; 7 AA.  
 AC P41875;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE FMRFAMIDE-LIKE NEUROPEPTIDE PF4 (KPNFIRFAMIDE).  
 OS Panagrellus redivivus.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 CC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RX MEDLINE=95232026; PubMed=7716079;  
 RA Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P.,  
 Thim L., Kubiak T.M., Martin R.A., Geary T.G.;  
 RT "Isolation and preliminary biological characterization of  
 RT KPNFIRFamide, a novel FMRFamide-related peptide from the free-living  
 RT nematode, Panagrellus redivivus.";  
 RL Peptides 16:87-93(1995).  
 CC -!- FUNCTION: MYOACTIVE; INDUCES A RAPID CONCENTRATION-DEPENDENT  
 CC MUSCLE TENSION INCREASE.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7  
 SQ SEQUENCE 7 AA; 921 MW; 69D40059C4576350 CRC64;

Query Match 25.0%; Score 13; DB 1; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+04;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PTF 4  
 :||  
 Db 2 PNF 4

## RESULT 13

AKH\_MELML  
 ID AKH\_MELML STANDARD; PRT; 8 AA.  
 AC P25423;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-FEB-1994 (Rel. 22, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE ADIPOKINETIC HORMONE (AKH).  
OS Melolontha melolontha (Cockchafer),  
OS Geotrupes stercorosus (Dor beetle), and  
OS Pachnoda marginata (Flower beetle).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
OC Scarabaeiformia; Scarabaeidae; Melolonthinae; Melolontha.  
RN [1]  
RN SEQUENCE.  
RC SPECIES=M. MELOLONTA, AND G. STERCOROSUS; TISSUE=CORPORA CARDIACA;  
RA MEDLINE=91248100; PubMed=2039445;  
RA Gaede G.;  
RT "A unique charged tyrosine-containing member of the adipokinetic  
RT hormone/red-pigment-concentrating hormone peptide family isolated and  
RT sequenced from two beetle species.";  
RT Biochem. J. 275:671-677(1991).  
RN [2]  
RN SEQUENCE.  
RC SPECIES=P. MARGINATA; TISSUE=CORPORA CARDIACA;  
RX MEDLINE=92265187; PubMed=1586453;  
RA Gaede G., Lopata A., Kellner R., Rinehart K.L. Jr.;  
RT "Primary structures of neuropeptides isolated from the corpora  
RT cardiaca of various cetonid beetle species determined by  
RT pulsed-liquid phase sequencing and tandem fast atom bombardment mass  
RT spectrometry.";  
RL Biol. Chem. Hoppe-Seyler 373:133-142(1992).  
CC -1- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA  
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF  
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT  
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.  
CC -1- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.  
DR PIR: S15422; S15422.  
DR PIR: S21663; S21663.  
DR INTERPRO: IPR002047; .  
DR PROSITE: PS00256; AKH; 1.  
DR MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 8 8 AMIDATION.  
FT MOD\_RES 8 8 AMIDATION.  
SQ SEQUENCE 8 AA; 1022 MW; 867AB775AB544736 CRC64;  
  
Query Match 25.0%; Score 13; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 8.9e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 7 NY 8  
Db 3 NY 4  
  
RESULT 14  
B44K\_PORGI ID B44K\_PORGI STANDARD; PRT; 8 AA.  
AC P81886;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 44 KDA IMMUNOGENIC PROTEIN (FRAGMENT).  
OS Porphyromonas gingivalis (Bacteroides gingivalis).  
OC Bacteria; CFB group; Bacteroidaceae; Porphyromonas.  
RN [1]  
RN SEQUENCE.  
RC STRAIN=VPB 3492;  
RX MEDLINE=20198497; PubMed=10731616;  
RA Norris J.M., Love D.N.;  
RT "Serum antibody responses of cats to soluble whole cell antigens of  
RT feline porphyromonas gingivalis.";  
RL Vet. Microbiol. 73:37-49(2000).  
CC -1- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.  
KW Antigen.  
FT NON\_TER 8 8  
SQ SEQUENCE 8 AA; 989 MW; 9554540326CB476D CRC64;

Query Match 25.0%; Score 13; DB 1; Length 8;  
Best Local Similarity 50.0%; Pred. No. 8.9e+04;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 2 PTFKEN 7  
Db 2 PYQKRN 7  
  
RESULT 15  
HTFL\_PERAM ID HTFL\_PERAM STANDARD; PRT; 8 AA.  
AC P04548;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE HYPERTREHALOSAEMIC FACTOR I (NEUROPEPTIDE M-I) (PERIPIANFTIN (C 1)  
DE (PEA-CAH-I) (LED-CC-I) (HYPERTREHALOSAEMIC NEUROPEPTIDE I).  
OS Periplaneta americana (American cockroach), and  
OS Leptinotarsa decemlineata (Colorado potato beetle), and  
OS Blatta orientalis (Oriental cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blattodea; Blattidae; Periplaneta.  
RN [1]  
RN SEQUENCE.  
RC SPECIES=P. AMERICANA;  
RX MEDLINE=85046530; PubMed=6548628;  
RA Witten J.L., Schaffer M.H., O'Shea M., Cook J.C., Hemling M.E.,  
RA Rinehart K.L. Jr.;  
RT "Structures of two cockroach neuropeptides assigned by fast atom  
RT bombardment mass spectrometry.";  
RL Biochem. Biophys. Res. Commun. 124:350-358(1984).  
RN [2]  
RN SEQUENCE.  
RC SPECIES=P. AMERICANA;  
RX MEDLINE=84298179; PubMed=6591205;  
RA Scarborough R.M., Jamieson G.C., Kalish F., Kramer S.J., McEnroe G.A.,  
RA Miller C.A., Schooley D.A.;  
RT "Isolation and primary structure of two peptides with  
RT cardioacceleratory and hyperglycemic activity from the corpora  
RT cardiaca of Periplaneta americana.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).  
RN [3]  
RN SEQUENCE.  
RC SPECIES=L. DECEMLINEATA; TISSUE=CORPORA CARDIACA;  
RX MEDLINE=90160053; PubMed=2576128;  
RA Gaede G., Kellner R.;  
RT "The metabolic neuropeptides of the corpus cardiaca from the potato  
RT beetle and the American cockroach are identical.";  
RL Peptides 10:1287-1289(1989).  
RN [4]  
RN SEQUENCE.  
RC SPECIES=B. ORIENTALIS; TISSUE=CORPORA CARDIACA;  
RX MEDLINE=90253659; PubMed=2340112;  
RA Gaede G., Rinehart K.L. Jr.;  
RT "Primary structures of hypertrehalosaemic neuropeptides isolated from  
RT the corpora cardiaca of the cockroaches Leucophaea maderae,  
RT Gromphadorhina portentosa, Blattella germanica and Blatta orientalis  
RT and of the stick insect Extatosoma tiaratum assigned by tandem fast  
RT atom bombardment mass spectrometry.";  
RL Biol. Chem. Hoppe-Seyler 371:345-354(1990).  
CC -1- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT  
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS  
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).  
CC -1- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.  
DR PIR: A05169; A05169.  
DR PIR: S08995; S08995.  
DR PIR: A49823; A49823.  
DR PIR: A44960; A44960.  
DR INTERPRO: IPR002047; .

DR PROSITE; PS00256; AKH; 1.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 8 8 AMIDATION.  
SQ SEQUENCE 8 AA; 991 MW; 86745775B9C452D6 CRC64;

Query Match 25.0%; Score 13; DB 1; Length 8;  
Best Local Similarity 40.0%; Pred. No. 8.9e+04;  
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 FKENY 8  
| |:  
Db 4 FSPNW 8

Search completed: January 17, 2001, 13:47:10  
Job time: 279 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:44:43 ; Search time 58.21 seconds  
(without alignments)  
18.122 Million cell updates/sec

Title: US-08-765-837-9  
Perfect score: 52  
Sequence: 1 DPTFKENYR 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 467

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_15:  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhmc:\*  
8: sp.organelle:\*  
9: sp.phase:\*  
10: sp.plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertibrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	42.3	8	Q9R5L7	Q9r5l7 clostridium
2	19	36.5	9	12 Q65711	Q65711 berne virus
3	18	34.6	8	10 Q9S8Z4	Q9s8z4 spinacia ol
4	17	32.7	9	12 Q82622	Q82622 avian infec
5	15	28.8	8	2 Q9R7T2	Q9r7t2 escherichia
6	14	26.9	7	8 Q98866	Q98866 spinacia ol
7	14	26.9	8	4 Q9V4J4	Q9v4j4 homo sapien
8	14	26.9	8	11 Q9QVJ8	Q9qvj8 mus sp. mep
9	14	26.9	9	4 Q9S574	Q9s574 homo sapien
10	13	25.0	7	5 Q9VYN9	Q9vyn9 drosophila
11	13	25.0	8	2 Q9R5R0	Q15898 homo sapien
12	13	25.0	8	4 Q15898	Q15898 shigella dy
13	13	25.0	8	4 Q9V4J3	Q9v4j3 homo sapien
14	13	25.0	8	8 Q36898	Q36898 nicotiana p
15	13	25.0	8	8 Q9TKES	Q9tkes leptospermu
16	13	25.0	8	12 Q9WJ33	Q9wj33 pseudomonas
17	13	25.0	9	2 Q51765	Q51765 pseudomonas
18	13	25.0	9	2 Q9R7H9	Q9r7h9 haemophilus
19	13	25.0	9	2 Q9R635	Q9r635 chlamydia t

20 13 25.0 9 2 Q9R5R1 Q9r5r1 shigella dy  
21 13 25.0 9 3 Q9P8E5 Q9p8e5 kluyveromyc  
22 13 25.0 9 4 Q9UMA0 Q9uma0 homo sapien  
23 13 25.0 9 8 Q316S3 Q316s3 anser caeru  
24 13 25.0 9 8 Q9TKG1 Q9tkg1 calothamnus  
25 13 25.0 9 8 Q9TKD9 Q9tkd9 pericalymma  
26 13 25.0 9 9 Q9XJN0 Q9xjn0 bacterioph  
27 13 25.0 9 12 Q64972 Q64972 avian rous-  
28 13 25.0 9 12 Q85723 Q85723 simlan sarc  
29 13 25.0 9 12 Q88612 Q88612 mlddelburq  
30 13 25.0 9 12 Q88953 Q88953 vaccinia vi  
31 12 23.1 8 2 Q52062 Q52062 bacillus me  
32 12 23.1 8 2 Q9R9C2 Q9r9c2 borrelia bu  
33 12 23.1 8 3 Q13591 Q13591 saccharomyc  
34 12 23.1 8 4 Q9UK84 Q9uk84 homo sapien  
35 12 23.1 8 4 Q9UJ50 Q9uj50 homo sapien  
36 12 23.1 8 10 Q42507 Q42507 triticum ae  
37 12 23.1 9 2 Q31363 Q31363 borrelia ga  
38 12 23.1 9 2 Q9R792 Q9r792 borrelia af  
39 12 23.1 9 2 Q9R3T0 Q9r3t0 borrelia af  
40 12 23.1 9 4 Q15999 Q15999 homo sapien  
41 12 23.1 9 4 Q9UKJ6 Q9ukj6 homo sapien  
42 12 23.1 9 8 Q9TLB0 Q9tld0 bostrychia  
43 12 23.1 9 8 Q9TKF2 Q9tkf2 asteromyrtu  
44 12 23.1 9 8 Q9TJ87 Q9tj87 caloglossa  
45 12 23.1 9 12 Q9IBY9 Q9iby9 kapost's sa

## ALIGNMENTS

RESULT 1

Q9R5L7 ID Q9R5L7 PRELIMINARY; PRT; 8 AA.  
AC Q9R5L7;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE 1,4-BETA-D-GLUCAN GLUCANOHYDROLASE (EC 3.2.1.4) (FRAGMENT).  
OS Clostridium thermocellum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1515;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92231850; PubMed=1567379;  
RA Romanec M.P., Fauth U., Kobayashi T., Huskisson N.S., Barker P.J.,  
RA Demain A.L.;  
RT "Purification and characterization of a new endoglucanase from  
RT Clostridium thermocellum";  
RL Biochem. J. 283:69-73(1992).  
SQ SEQUENCE 8 AA; 823 MW; C2CIAB1DD9D1B775 CRC64;

Query Match 42.3%; Score 22; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred.No. 3.7e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PTFKE 6

Db 2 PTFKE 6

RESULT 2

Q65711 ID Q65711 PRELIMINARY; PRT; 9 AA.  
AC Q65711;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE ORF3 (FRAGMENT).  
OS Berne virus (BEV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;

OC Coronaviridae; Torovirus.  
 OX NCBI\_TaxID=111156;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=50080137; PubMed=2293666;  
 RA Snijder E.J., Horzinek M.C., Spaan W.J.M.;  
 RT "A 3'-coterminally nested set of independently transcribed mRNAs is  
 generated during Berne virus replication.";  
 RL J. Virol. 64:331-338(1990).  
 DR EMBL: M33502; AAA42817.1; -;  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1234 MW; D8EE736B5451AB19 CRC64;

Query Match 36.5%; Score 19; DB 12; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 FKENY 8  
 I : I I  
 Db 2 FETNY 6

## RESULT 3

ID Q9S824 PRELIMINARY; PRT; 8 AA.  
 AC Q9S824;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE PHOTOSYSTEM I REACTION CENTER SUBUNIT IV, PSI-E.  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;  
 OC Caryophyllales; Chenopodiaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92249324; PubMed=1374333;  
 RA Lagoutte B., Vallon O.;  
 RT "Purification and membrane topology of PSI-D and PSI-E, two subunits  
 of the photosystem I reaction center.";  
 RL Eur. J. Biochem. 205:1175-1185(1992).  
 SQ SEQUENCE 8 AA; 1082 MW; 2145BBI324069044 CRC64;

Query Match 34.6%; Score 18; DB 10; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 KENY 8  
 I : I I  
 Db 5 KESY 8

## RESULT 4

ID Q82622 PRELIMINARY; PRT; 9 AA.  
 AC Q82622;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE ORF A (FRAGMENT).  
 OS avian infectious bronchitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=111120;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86010264; PubMed=2995560;  
 RA Bourns M.E., Binns M.M., Brown T.D.;  
 RT "Sequencing of coronavirus IBV genomic RNA: three open reading frames  
 in the 5' 'unique' region of mRNA D.";

RL J. Gen. Virol. 66:2253-2258(1985).  
 DR EMBL: M29338; AAA46234.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 1134 MW; D2C4B3327741B46 CRC64;

Query Match 32.7%; Score 17; DB 12; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ENYR 9  
 I : I I  
 Db 1 EOYR 4

## RESULT 5

ID Q9R7T2 PRELIMINARY; PRT; 8 AA.  
 AC Q9R7T2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE HYPOTHETICAL PROTEIN HI0064 (FRAGMENT).  
 GN YQFG.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding  
 to the 12.7-28.0 min Region on the Linkage Map.";  
 RL DNA Res. 3:137-155(1996).  
 DR EMBL: D90705; BAA35310.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 964 MW; DF133B1DD04B476A CRC64;

Query Match 28.8%; Score 15; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DPTFK 6  
 I : I I  
 Db 1 DPYIAE 6

## RESULT 6

ID O98866 PRELIMINARY; PRT; 7 AA.  
 AC O98866;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE CYTOCHROME B/F SUBUNIT IV (FRAGMENT).  
 OS Spinacia oleracea (Spinach).  
 OC Chloroplast.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;  
 OC Caryophyllales; Chenopodiaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86120353; PubMed=3003688;  
 RA Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;  
 RT "Spinach plastid genes coding for initiation factor IF-1, ribosomal



RT protein S11 and RNA polymerase alpha-subunit.";  
 RL Nucleic Acids Res. 14:1029-1044(1986).  
 DR EMBL: X03496; CAA27215.1; -.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 907 MW; 644729D77409C420 CRC64;

Query Match 26.9%; Score 14; DB 8; Length 7;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 NYR 9  
 Db 1 NFR 3

RESULT 7  
 ID Q9Y4J4 PRELIMINARY; PRT; 8 AA.  
 AC Q9Y4J4;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
 DE RUNT/82NT/MTG8 PROTEIN (FRAGMENT).  
 GN RUNT/82NT/MTG8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95002916; PubMed=7919324;  
 RA Tighe J.E., Calabi F.;  
 RT "Alternative, out-of-frame runt/MTG8 transcripts are encoded by the  
 RT derivative (8) chromosome in the t(8;21) of acute myeloid leukemia  
 RT M2.";  
 RL Blood 84:2115-2121(1994).  
 DR EMBL: S74092; AAD14144.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 8 AA; 1067 MW; 20F414044B17244B CRC64;

Query Match 26.9%; Score 14; DB 4; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 ENYR 9  
 Db 4 ENRR 7

RESULT 8  
 ID Q9QVJ8 PRELIMINARY; PRT; 8 AA.  
 AC Q9QVJ8;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE MEPRIN-A.  
 OS Mus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10095;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=91373354; PubMed=1894522;  
 RA Kounnas M.Z., Wolz R.L., Gorbea C.M., Bond J.S.;  
 RT "Meprin-A and -B. Cell surface endopeptidases of the mouse kidney.";  
 RL J. Biol. Chem. 266:17350-17357(1991).  
 SQ SEQUENCE 8 AA; 877 MW; 43A5A76AB4069DD4 CRC64;

Query Match 26.9%; Score 14; DB 11; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPT 3  
 Db 5 DPS 7

RESULT 9  
 ID O95574 PRELIMINARY; PRT; 9 AA.  
 AC O95574;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE SERUM ALBUMIN (FRAGMENT).  
 GN ALB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=FETAL LIVER;  
 MEDLINE=82081882; PubMed=6171778;  
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
 RA Najarian R.C., Seeburg P.H., Wion K.L.;  
 RT "The sequence of human serum albumin cDNA and its expression in E.  
 RT coli.";  
 RL Nucleic Acids Res. 9:6103-6114(1981).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=FETAL LIVER;  
 MEDLINE=82105994; PubMed=6275391;  
 RA Dugaiczky A., Law S.W., Dennison O.E.;  
 RT "Nucleotide sequence and the encoded amino acids of human serum  
 RT albumin mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=FETAL LIVER;  
 MEDLINE=85155492; PubMed=6085063;  
 RA Urano Y., Sakai M., Watanabe K., Tamaoki T.;  
 RT "Random arrangement of the albumin and alpha-fetoprotein genes in the  
 RT human genome.";  
 RL Gene 32:255-261(1984).  
 DR EMBL: L00133; AAC95394.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 9 AA; 1186 MW; 92340B0400440681 CRC64;

Query Match 26.9%; Score 14; DB 4; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PTFK 5  
 Db 1 PTMR 4

RESULT 10  
 ID Q9VYN9 PRELIMINARY; PRT; 7 AA.  
 AC Q9VYN9;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE CG18256 PROTEIN (FRAGMENT).  
 GN CG18256.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscophora;

OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballieu R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hartis N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003488; AAF48153.1; -;  
 DR FLYBASE; FBgn0030380; CG18256.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 817 MW; 69C737644731A740 CRC64;

Query Match 25.0%; Score 13; DB 5; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DPTF 4  
 Db :| |  
 4 NPLF 7  
 RESULT 11  
 Q9R5R0 PRELIMINARY; PRT; 8 AA.  
 AC Q9R5R0;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DE 11,500 DA PRODUCT OF ORFA.  
 OS Shigella dysenteriae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OX Shigella.  
 OX NCBI\_TaxID=622;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=92085268; PubMed=1660923;  
 RA Polard P., Prere M.F., Chandler M., Fayet O.;  
 RT "Programmed translational frameshifting and initiation at an AUU codon  
 in gene expression of bacterial insertion sequence IS911.";  
 RL J. Mol. Biol. 222:465-477(1991).  
 SQ SEQUENCE 8 AA; 935 MW; F4C5B9C45AE33336 CRC64;  
 Query Match 25.0%; Score 13; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 KENY 8  
 Db :| |  
 3 KXNF 6  
 RESULT 12  
 Q15898 PRELIMINARY; PRT; 8 AA.  
 ID Q15898;  
 AC Q15898;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DE 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
 DE (CLONE XP6A11B) (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Lee C.C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,  
 RA Coolbaugh M.I., Chnault C.A., Baldini A., Lindsay E.A., Zhao Z.Y.,  
 RA Caskey C.T.H.;  
 RL Hum. Mol. Genet. 0:0-0(0).  
 DR EMBL; L32078; AAA73888.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 938 MW; 34A415B0477B45BB CRC64;  
 Query Match 25.0%; Score 13; DB 4; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 ENY 8  
 Db :| |  
 1 ESY 3  
 RESULT 13  
 Q9Y4J3 PRELIMINARY; PRT; 8 AA.  
 ID Q9Y4J3;  
 AC Q9Y4J3;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
 DE 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
 DE RUNT/68NT/MTG8 PROTEIN (FRAGMENT).  
 GN RUNT/68NT/MTG8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BONE MARROW;  
 RX MEDLINE=95002916; PubMed=7919324;  
 RA Tighe J.E., Calabi F.;  
 RT "Alternative, out-of-frame runt/MTG8 transcripts are encoded by the  
 derivative (8) chromosome in the t(8;21) of acute myeloid leukemia  
 M2.";  
 RL Blood 84:2115-2121(1994).

DR EMBL; S74094; AAD14973.2; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 8 AA; 929 MW; 308764405817244B CRC64;

Query Match 25.0%; Score 13; DB 4; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPT 3  
 :||  
 Db 6 NPT 8

## RESULT 14

Q36898 PRELIMINARY; PRT; 8 AA.  
 AC Q36898;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE RPS19; PROTEIN (FRAGMENT).  
 GN RPS19.  
 OS Nicotiana glauca, and Nicotiana glauca (Bigelow's tobacco).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asterids I;  
 OC Solanales; Solanaceae; Nicotiana.  
 ON NCBI\_TaxID=49453, 4088;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Goulding S.E., Olmstead R.G., Morden C.W., Wolfe K.H.;  
 RL Mol. Gen. Genet. 0:0-0(1996).  
 DR EMBL; Z71234; CAA94933.1; -.  
 DR EMBL; Z71225; CAA94921.1; -.  
 KW Chloroplast.  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 977 MW; FD43333735A411A6 CRC64;

Query Match 25.0%; Score 13; DB 8; Length 8;  
 Best Local Similarity 40.0%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TFEN 7  
 :||  
 Db 4 SLKN 8

## RESULT 15

Q9TKES PRELIMINARY; PRT; 8 AA.  
 AC Q9TKES;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE ATPB (FRAGMENT).  
 GN ATPB.  
 OS Leptospermum erubescens.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Myrtales; Myrtaceae; Leptospermum.  
 ON NCBI\_TaxID=106049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA O'Brien M.M., Quinn C.J., Wilson P.G.;  
 RT "Molecular Systematics of the Leptospermum Suballiance (Myrtaceae).";  
 RL Aust. J. Bot. 48:0-0(2000).  
 DR EMBL; AF184690; AAF03860.1; -.  
 KW Chloroplast.  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 876 MW; ECA1B1B764405056 CRC64;

Query Match 25.0%; Score 13; DB 8; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 3.7e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 DPT 3  
 :||  
 Db 4 NPT 6

Search completed: January 17, 2001, 13:44:44  
 Job time: 198 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2001, 13:42:58 ; Search time 31.16 Seconds  
(without alignments)  
5.187 Million cell updates/sec

Title: US-08-765-837-9  
Sequence: 1 DPTFKENYR 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 44977

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	55.8	9	1	US-08-077-256-4
2	29	55.8	9	1	US-08-259-672-4
3	29	55.8	9	1	US-08-459-351-4
4	29	55.8	9	1	US-08-460-533-4
5	29	55.8	9	3	US-08-814-836-3
6	29	55.8	9	4	PCT-US94-06654-4
7	24	46.2	5	1	US-07-926-786-1
8	23	44.2	7	1	US-08-325-553-5
9	23	44.2	7	2	US-08-394-152A-5
10	23	44.2	8	2	US-08-669-284B-35
11	23	44.2	8	2	US-08-310-912A-65
12	23	44.2	8	3	US-08-841-089-65
13	23	44.2	8	4	PCT-US95-04570-65
14	23	44.2	8	4	PCT-US95-04589-65
15	22	42.3	8	1	US-08-094-948A-21
16	22	42.3	8	4	PCT-US96-09319-21
17	22	42.3	9	1	US-08-360-706-3
18	22	42.3	9	1	US-08-360-706-7
19	22	42.3	9	1	US-08-467-083-26
20	22	42.3	9	1	US-08-414-417B-26
21	22	42.3	9	2	US-08-486-348A-26
22	22	42.3	9	3	US-08-468-545B-26
23	22	42.3	9	3	US-08-159-339A-256
24	22	42.3	9	3	US-08-466-680B-26
25	21	40.4	8	2	US-08-891-848-11
26	21	40.4	9	1	US-08-261-206A-51
27	21	40.4	9	2	US-08-416-870C-22
28	21	40.4	9	5	5496924-38

29	20	38.5	5	1	US-07-926-786-2	Sequence 2, Appli
30	20	38.5	5	1	US-08-361-708-7	Sequence 7, Appli
31	20	38.5	5	1	US-08-536-277-7	Sequence 7, Appli
32	20	38.5	6	1	US-08-136-743B-67	Sequence 67, Appli
33	20	38.5	6	3	US-09-040-216-58	Sequence 58, Appli
34	20	38.5	7	1	US-08-136-743B-66	Sequence 66, Appli
35	20	38.5	7	2	US-08-540-406-13	Sequence 13, Appli
36	20	38.5	7	2	US-08-680-326-102	Sequence 102, Appli
37	20	38.5	7	3	US-08-656-055-13	Sequence 13, Appli
38	20	38.5	7	3	US-09-040-216-7	Sequence 7, Appli
39	20	38.5	7	3	US-09-040-216-57	Sequence 57, Appli
40	20	38.5	7	4	PCT-US95-13233-13	Sequence 13, Appli
41	20	38.5	7	5	5194585-11	Patent No. 5194585
42	20	38.5	8	3	US-09-045-632-102	Sequence 102, Appli
43	20	38.5	8	3	US-08-722-258-69	Sequence 69, Appli
44	20	38.5	8	3	US-08-159-339A-1123	Sequence 1123, Ap
45	20	38.5	9	1	US-08-426-819A-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1  
US-08-077-256-4  
; Sequence 4, Application US/08077256  
; Patent No. 5582995  
; GENERAL INFORMATION:  
; APPLICANT: Joseph Avruch, M.D.  
; APPLICANT: Xian-feng Zhang, Ph.D.  
; TITLE OF INVENTION: INHIBITING PROTEIN  
; TITLE OF INVENTION: INTERACTIONS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA: US/08/077,256  
; APPLICATION NUMBER: US/08/077,256  
; FILING DATE: 19930611  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark, Esq.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/190001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-077-256-4

Query Match 55.8%; Score 29; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 DPTFKENY 8

Db 2 DPTIEDSY 9  
||| :||

## RESULT 2

US-08-259-672-4  
; Sequence 4, Application US/08259672  
; Patent No. 5736337  
; GENERAL INFORMATION:  
; APPLICANT: Joseph Avruch  
; APPLICANT: Xian-feng Zhang  
; APPLICANT: Mark S. Marshall  
; TITLE OF INVENTION: INHIBITING PROTEIN  
; INTERACTIONS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55sx  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/259,672  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/077,256  
; FILING DATE: June 11, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark, Esq.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/234001  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-259-672-4

Query Match 55.8%; Score 29; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFKENY 8  
||| :||  
Db 2 DPTIEDSY 9

## RESULT 3

US-08-459-351-4  
; Sequence 4, Application US/08459351  
; Patent No. 5763571  
; GENERAL INFORMATION:  
; APPLICANT: Joseph Avruch  
; APPLICANT: Xian-feng Zhang  
; APPLICANT: Mark S. Marshall  
; TITLE OF INVENTION: INHIBITING PROTEIN INTERACTIONS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street

; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55sx  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,351  
; FILING DATE: June 2, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/259,672  
; FILING DATE: June 10, 1994  
; APPLICATION NUMBER: 08/077,256  
; FILING DATE: June 11, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark, Esq.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/234003  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-459-351-4

Query Match 55.8%; Score 29; DB 1; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFKENY 8  
||| :||  
Db 2 DPTIEDSY 9

## RESULT 4

US-08-460-533-4  
; Sequence 4, Application US/08460533  
; Patent No. 5767075  
; GENERAL INFORMATION:  
; APPLICANT: Joseph Avruch  
; APPLICANT: Xian-feng Zhang  
; APPLICANT: Mark S. Marshall  
; TITLE OF INVENTION: INHIBITING PROTEIN INTERACTIONS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55sx  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,533  
; FILING DATE: June 2, 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/259,672  
; FILING DATE: June 10, 1994

APPLICATION NUMBER: 08/077,256  
FILING DATE: June 11, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul T. Clark, Esq.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/234002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-460-533-4

Query Match 55.8% Score 29; DB 1; Length 9;  
Best Local Similarity 50.08; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0;

Qy 1 DPTFKENY 8  
||| : : :  
Db 2 DPTIEDSY 9

RESULT 5  
US-08-814-836-3  
Sequence 3, Application US/08814836  
Patent No. 6103692  
GENERAL INFORMATION:  
APPLICANT: Avruch, Joseph  
APPLICANT: Luo, ZhuJun  
APPLICANT: Marshall, Mark S.  
TITLE OF INVENTION: INHIBITING PROTEIN  
TITLE OF INVENTION: INTERACTIONS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/814,836  
FILING DATE: 11-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/013,274  
FILING DATE: 12-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 00786/313001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6103692e  
US-08-814-836-3

Query Match 55.8% Score 29; DB 3; Length 9;  
Best Local Similarity 50.08; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0;

Qy 1 DPTFKENY 8  
||| : : :  
Db 2 DPTIEDSY 9

RESULT 6  
PCT-US94-06654-4  
Sequence 4, Application PC/TUS9406654  
GENERAL INFORMATION:  
APPLICANT: Joseph Avruch  
APPLICANT: Xian-feng Zhang  
APPLICANT: Mark S. Marshall  
TITLE OF INVENTION: INHIBITING PROTEIN INTERACTIONS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/06654  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul T. Clark, Esq.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/234001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
PCT-US94-06654-4

Query Match 55.8% Score 29; DB 4; Length 9;  
Best Local Similarity 50.08; Pred. No. 1.3e+05;  
Matches 4; Conservative 3; Mismatches 1; Indels 0;

Qy 1 DPTFKENY 8  
||| : : :  
Db 2 DPTIEDSY 9

RESULT 7  
US-07-926-786-1  
Sequence 1, Application US/07926786  
Patent No. 5280113  
GENERAL INFORMATION:  
APPLICANT: Rademacher, Thomas W.  
APPLICANT: Manger, Ian D.  
APPLICANT: Wong, Simon  
APPLICANT: Dwek, Raymond A.

;; TITLE OF INVENTION: Method for Producing Synthetic N-Linked  
;; TITLE OF INVENTION: Glycoconjugates  
;; NUMBER OF SEQUENCES: 3  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SG  
;; STREET: 800 N. Lindbergh Blvd.  
;; CITY: St. Louis  
;; STATE: Missouri  
;; COUNTRY: USA  
;; ZIP: 63167  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/926,786  
;; FILING DATE: 19920811  
;; CLASSIFICATION: 536  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/776911  
;; FILING DATE: 15-OCT-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/394691  
;; FILING DATE: 16-AUG-1989  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Meyer, Scott J.  
;; REGISTRATION NUMBER: 25,275  
;; REFERENCE/DOCKET NUMBER: 07-27(904)A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (314)694-3117  
;; TELEFAX: (314)694-5435  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-07-926-786-1

Query Match 46.2%; Score 24; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPTF 4  
Db 2 DPTF 5

RESULT 8  
US-08-325-553-5  
; Sequence 5, Application US/08325553  
; Patent No. 553866  
; GENERAL INFORMATION:  
; APPLICANT: Israeli, Ron S.  
; APPLICANT: Heston, Warren D.W.  
; APPLICANT: Fair, William R.  
; TITLE OF INVENTION: THE PROSTATE-SPECIFIC MEMBRANE ANTIGEN  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/325,553  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/973,337A  
;; FILING DATE: 05 NOV 1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: White, John P.  
;; REGISTRATION NUMBER: 28,678  
;; REFERENCE/DOCKET NUMBER: 1747/41426  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 977-9550  
;; TELEFAX: (212) 664-0525  
;; TELEX: 422523 COOP UI  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapien  
;; TISSUE TYPE: Carcinoma  
;; IMMEDIATE SOURCE:  
;; CLONE: Prostate Specific Membrane Antigen  
US-08-325-553-5  
Query Match 44.2%; Score 23; DB 1; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 DPTF 5  
Db 3 DPTF 7  
RESULT 9  
US-08-394-152A-5  
; Sequence 5, Application US/08394152A  
; Patent No. 5935818  
; GENERAL INFORMATION:  
; APPLICANT: Israeli, Ron S.  
; APPLICANT: Heston, Warren D.W.  
; APPLICANT: Fair, William R.  
; TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM 330 466 DX2  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,152A  
; FILING DATE: 24-FEB-95  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 41426-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400



TELEPHONE: (212) 391-0525  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Carcinoma  
IMMEDIATE SOURCE:  
CLONE: Prostate Specific Membrane Antigen  
US-08-394-152A-5

Query Match 44.2%; Score 23; DB 2; Length 7;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPTFK 5  
|| ||  
Db 3 DPMFK 7

RESULT 10  
US-08-669-284B-35  
; Sequence 35, Application US/08669284B  
; Patent No. 5939534  
; GENERAL INFORMATION:  
; APPLICANT: Inoue, Makoto  
; APPLICANT: Kikuchi, Kaoru  
; APPLICANT: Ishige, Yoko  
; APPLICANT: Ito, Akira  
; APPLICANT: Kimura, Toru  
; APPLICANT: Nakayama, Chikao  
; APPLICANT: No. 5939534uchi, Hiroshi  
; TITLE OF INVENTION: NOVEL HUMAN CILIARY NEUTROPHILIC FACTORS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/669,284B  
; FILING DATE: 28-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP94/02269  
; FILING DATE: 27-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 06-268281  
; FILING DATE: 05-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 06-201504  
; FILING DATE: 02-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 05-350934  
; FILING DATE: 29-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Nakamura, Dean H.  
; REGISTRATION NUMBER: 33,981  
; REFERENCE/DOCKET NUMBER: Q-42041  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-669-284B-35

Query Match 44.2%; Score 23; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPKEN 7  
:||||  
Db 4 SPKEN 8

RESULT 11  
US-08-310-912A-65  
; Sequence 65, Application US/08310912A  
; Patent No. 5981730  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindrinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION  
; NUMBER OF SEQUENCES: 208  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2904  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/310,912A  
; FILING DATE: September 22, 1994  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/227,360  
; FILING DATE: April 13, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lech, Karen F.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/254001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 100254  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-310-912A-65

Query Match 44.2%; Score 23; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PTFKE 6  
   | | | |  
Db 3 PTFSE 7

## RESULT 12

US-08-841-089-65  
; Sequence 65, Application US/08841089  
; Patent No. 6127607  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindrinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street Suite 3100  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2904  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/841,089  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/227,360  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/230001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 100254  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-841-089-65

Query Match 44.2%; Score 23; DB 3; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PTFKE 6  
   | | | |  
Db 3 PTFSE 7

## RESULT 13

PCT-US95-04570-65

; Sequence 65, Application PC/TUS9504570  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindrinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street Suite 3100  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2904  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/04570  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/227,360  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/230001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 100254  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-04570-65

Query Match 44.2%; Score 23; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PTFKE 6  
   | | | |  
Db 3 PTFSE 7

## RESULT 14

PCT-US95-04589-65  
; Sequence 65, Application PC/TUS9504589  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindrinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 201  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2904  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04589  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/227,360  
FILING DATE: 13-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/230001  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 100254  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-04589-65

Query Match 44.2%; Score 23; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.3e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PTFKE 6  
|||  
Db 3 PTFSE 7

RESULT 15  
US-08-094-948A-21  
Sequence 21, Application US/08094948A  
Patent No. 5621075  
GENERAL INFORMATION:  
APPLICANT: Kahn, C. Ronald  
APPLICANT: White, Morris F.  
APPLICANT: Rothenberg, Paul Louis  
TITLE OF INVENTION: INSULIN RECEPTOR SUBSTRATE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lahive & Cockfield  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/094,948A  
FILING DATE: 21-JULY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/643,982  
FILING DATE: 18-JAN-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis (PLM)  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: JDP-013DV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-094-948A-21

Query Match 42.3%; Score 22; DB 1; Length 8;  
Best Local Similarity 42.9%; Pred. No. 1.3e+05;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TFKENYR 9  
|||  
Db 1 TFEBSFQ 7

Search completed: January 17, 2001, 13:42:59  
Job time: 158 sec

